PCT/EP 98/06971 #7

10.02.99

09/529739

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Patent Office
Canberra

REC'D	12 FEB 1999
WIPO	PCT

I, KIM MARSHALL, MANAGER EXAMINATION SUPPORT AND SALES, hereby certify that the annexed is a true copy of the Provisional specification in connection with Application No. PO 9745 for a patent by RHONE-POULENC AGROCHIMIE filed on 10 October 1997.

WITNESS my hand this Fifth day of January 1999

KIM MARSHALL
MANAGER EXAMINATION SUPPORT AND SALES

H 03.02.99

S & F Ref: 395498

ORIGINAL

AUSTRALIA

Patents Act 1990

PROVISIONAL SPECIFICATION FOR THE INVENTION ENTITLED:

Methods for Obtaining Plant Varieties

Name and Address of Applicant: Rhone-Poulenc Agrochimie, of 14-20, rue Pierre Baizet, BP 9163, 69263 Lyon Cedex 09, FRANCE

This invention is best described in the following statement:

Methods for Obtaining Plant Varieties

TECHNICAL FIELD

The present invention relates to nucleotide sequences which encode polypeptides involved in the DNA mismatch repair systems of plants, and to the polypeptides encoded by those nucleotide sequences. The invention also relates to nucleotide sequences and polypeptide sequences for use in altering the DNA mismatch repair system in plants. The invention also relates to a process for altering the DNA mismatch repair system of a plant cell, to a process for increasing genetic variations in plants and to processes for obtaining plants having a desired characteristic.

BACKGROUND OF THE INVENTION

Plant breeding essentially relies on and makes use of genetic variation which occurs naturally within and between members of a family, a genus, a species or a subspecies. Another source of genetic variation is the introduction of genes from other organisms which may or may not be related to the host plant.

Allelic loci or non-allelic genes which constitute or contribute to desired quantitative (e.g. growth performance, yield, etc.) or qualitative (e.g. deposition, content and composition of seed storage products; pathogen resistance genes; etc.) traits that are absent, incomplete or inefficient in a species or subspecies of interest are typically introduced by the plant breeder from other species or subspecies, or *de novo*. This introduction is often done by crossing, provided that the species to be crossed are sexually compatible. Other means of introducing genomes, individual chromosomes or genes into plant cells or plants are well known in the art. They include cell fusion, chemically aided transfection (Schocher et al., 1986, Biotechnology 4: 1093) and ballistic (McCabe et al., 1988, Biotechnology 6: 923), microinjection (Neuhaus et al., 1987, TAG 75: 30), electroporation of protoplasts (Chupeau et al., 1989, Biotechnology 7: 53) or microbial transformation methods such as Agrobacterium mediated transformation (Horsch et al., 1985, Science 227: 1229; Hiei et al., 1996, Biotechnology 14: 745).

However, when a foreign genome, chromosome or gene is introduced into a plant, it will often segregate in subsequent generations from the genome of the recipient plant or plant cell during mitotic and meiotic cell divisions and, in consequence, become lost from the host plant or plant cell into which it had been introduced. Occasionally, however, the introduced genome, chromosome or gene physically combines entirely or in part with the genome, chromosome or gene of the host plant or plant cell in a process which is called recombination.

Recombination involves the exchange of covalent linkages between DNA molecules in regions of identical or similar sequence. It is referred to here as homologous recombination if donor and recipient DNA are identical or nearly identical (at least 98%

base sequence identity), and as homeologous recombination if donor and recipient DNA are not identical but are similar (less than 98% base sequence identity).

The ability of two genomes, chromosomes or genes to recombine is known to depend largely on the evolutionary relation between them and thus on the degree of sequence similarity between the two DNA molecules. Whereas homologous recombination is frequently observed during mitosis and meiosis, homeologous recombination is rarely or never seen.

From a breeder's perspective, the limits within which homologous recombination occurs, therefore, define a genetic barrier between species, varieties or lines, in contrast to homeologous recombination which can break this barrier. Homeologous recombination is thus of great importance for plant breeding. Accordingly there is a need for a process for enhancing the frequency of homeologous recombination in plants. In particular, there is a need for a process of increasing homeologous recombination to significantly shorten the length of breeding programs by reducing the number of crosses required to obtain an otherwise rare recombination event.

At least in *Escherichia coli*, homologous and homeologous recombination are known to share a common pathway that requires among others the proteins RecA, RecB, RecC, RecD and makes use of the SOS induced RuvA and RuvB, respectively. It has been suggested that mating induced recombination follows the Double-Strand Break Repair model (Szostak et al., 1983, Cell 33, 25-35), which is widely used to describe genetic recombination in eukaryotes. Following the alignment of homologous or homeologous DNA double helices the RecA protein mediates an exchange of a single DNA strand from the donor helix to the aligned recipient DNA helix. The incoming strand screens the recipient helix for sequence complementarity, seeking to form a heteroduplex by hydrogen bonding the complementary strand. The displaced homologous or homeologous strand of the recipient helix is guided into the donor helix where it base pairs with its counterpart strand to form a second heteroduplex. The resulting branch point then migrates along the aligned chromosomes thereby elongating and thus stabilising the initial heteroduplexes. Single stranded gaps (if present) are closed by DNA synthesis. The strand cross overs (Holliday junction) are eventually resolved enzymatically to yield the recombination products.

Although in wild-type *E. coli* homologous and homeologous recombination are thus mechanistically similar if not identical, homologous recombination in conjugational crosses *E. coli* x *E. coli* occurs five orders of magnitude more frequently than homeologous recombination in conjugational crosses *E. coli* x *S. typhimurium* (Matic et al. 1995; Cell 80, 507-515). The imbalance in favour of homologous recombination was shown to be caused largely by the bacterial MisMatch Repair (MMR) system since its

inactivation increased the frequency of homeologous recombination in *E. coli* up to 1000 fold (Rayssiguier et al. 1989, Nature 342, 396-401).

In *E. coli*, the MMR system (reviewed by Modrich 1991, Annual Rev Genetics 25, 229-253) is composed of only three proteins known as MutS, MutL and MutH. MutS recognizes and binds to base pair mismatches. MutL then forms a stable complex with mismatch bound MutS. This protein complex now activates the MutH intrinsic single stranded endonuclease which nicks the strand containing the misplaced base and thereby prepares the template for DNA repair enzymes.

During recombination, MMR components inhibit homeologous recombination. In vitro experiments demonstrated that MutS in complex with MutL binds to mismatches at the recombination branch point and physically blocks RecA mediated strand exchange and heteroduplex formation (Worth et al., 1994; PNAS 91, 3238-3241). Interestingly, the SOS dependent RuvAB mediated branch migration is insensitive to MutS/MutL, explaining the observed slight increase in SOS dependent homeologous recombination. Homeologous mating even induces the SOS response, thereby taking advantage of RuvAB induction (Matic et al. 1995, Cell 80, 507-515).

The MMR system thus appears to be a genetic guardian over genome stability in *E. coli*. In this role it essentially determines the extent to which genetic isolation, that is, speciation, occurs. The diminished sensitivity of the SOS system to MMR, however, allows (within limits) for rapid genomic changes at times of stress, providing the means for fast adaptation to altered environmental conditions and thus contributing to intraspecies genetic variation and species evolution.

The important role of MMR in preserving genomic integrity has been established also in certain eukaryotes. In its efficiency, the human MMR, for example, may even counteract potential gene therapy tools such as triple-helix forming oligonucleotides including RNA-DNA hybrid molecules (Havre et al., 1993, J. Virology 67: 7234-7331; Wang et al., 1995, Mol. Cell. Biol. 15: 1759-1768; Kotani et al., 1996, Mol. Gen. Genetics 250: 626-634; Cole-Strauss et al., 1996, Science 273: 1387-1389). Such oligonucleotides are designed to introduce single base changes into selected DNA target sequences in order to inactivate for example cancer genes or to restore their normal function. The resulting base mismatches however are recognised by the mismatch repair system which then directs removal of the mismatched base, thereby reducing the efficiency of oligonucleotide induced site-specific mutagenesis.

To date, two families of related genes, homologous to the bacterial *Mut*S and *Mut*L genes have been identified or isolated in yeast and mammals (recent reviews by Arnheim and Shibata, 1997, Curr. Opinion Genet. Dev. 7, 364-370; Modrich and Lahue, 1996, Annual Rev. Biochem. 65, 101-133; Umar and Kunkel, 1996, Eur. J. Biochem. 238, 297-307). Biochemical and genetic analysis indicated that eukaryotic MutS homologs

(MSH) and MutL homologs (MLH, PMS), respectively, fulfil similar protein functions as their bacterial counterparts. Their relative abundance, however, could reflect different mismatch specificity and/or specialisation for different tissues or organelles or developmental processes such as mitotic versus meiotic recombination.

To date, six different genes homologous to *Mut*S have been isolated in yeast (y*MSH*), and their homologs have been found in mouse (m*MSH*) and human (h*MSH*), respectively. Encoded proteins yMSH2, yMSH3 and yMSH6 appear to be the main MutS homologs involved in MMR during mitosis and meiosis in yeast, where the complementary proteins MSH3 and MSH6 alternatively associate with MSH2 to recognise different mismatch substrates (Masischky et al., 1996, Genes Dev. 10, 407-420). Similar protein interactions have been demonstrated for the human homologs hMSH2, hMSH3 and hMSH6 (Acharya et al., 1996, PNAS 93, 13629-13634).

MutL homologs (MLH and PMS), recently reviewed by Modrich and Lahue (1996, Annual Rev. Biochem. 65, 101-133) have so far been found in yeast (yMLH1 and yPMS1), mouse (mPMS2) and human (hMLH1, hPMS1 and hPMS2). The h*PMS*2 is a member of a family of at least 7 genes (Horii et al., 1994, Biochem. Biophys. Res. Commun. 204, 1257-1264) and its gene product is most closely related to yPMS1. Prolla et al. (1994, Science 265, 1091-1093) presented evidence for yPMS1 and yMLH1 to physically associate with each other and, together, to interact with the MutS homolog yMSH2 to form a ternary complex involved in mismatch substrate binding.

However, while medical interest in mismatch repair has prompted extensive research on MMR in bacteria, yeast and mammals, MMR genes have not been isolated from higher plants prior to the present invention and no attempts to adjust the plant MMR to plant breeding needs have been reported.

SUMMARY OF THE INVENTION

According to a first embodiment of the invention, there is provided an isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide involved in the DNA mismatch repair system of a plant. In one form of this embodiment, the invention provides an isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human. More particularly, the invention provides the coding sequences of the genes *AtMSH3* and *AtMSH6* of *Arabidopsis thaliana*, as defined hereinbelow.

According to a second embodiment of the invention, there is provided a polypeptide functionally involved in the DNA mismatch repair system of a plant, for example a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a

human such as a polypeptide encoded by the genes *AtMSH3* or *AtMSH6* of *Arabidopsis thaliana*, as defined hereinbelow.

According to a third embodiment of the invention, there is provided an isolated and purified DNA molecule comprising a polynucleotide sequence selected from the group consisting of (i) a sequence encoding a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence; and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant.

According to a fourth embodiment of the invention there is provided a chimeric gene comprising a DNA sequence selected from the group consisting of (i) a sequence encoding a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence; and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant, together with regulation elements capable of functioning in a plant cell. Examples of such regulation elements include constitutive, inducible, tissue type specific and cell type specific promoters such as 35S, NOS, PR1a, AoPR1 and DMC1.

In the third and fourth embodiments, said interference, by said polynucleotide sequence, with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair peptide of a yeast or a human typically occurs by hybridisation or by co-suppression.

According to a fifth embodiment of the invention there is provided a plasmid or vector comprising a chimeric gene of the fourth embodiment. A vector of the fifth embodiment may be, for example, a viral vector.

According to a sixth embodiment of the invention, there is provided a plant cell transformed, transfected or electroporated with a plasmid or vector of the fifth embodiment.

According to seventh embodiment of the invention, there is provided a plant comprising a cell of the sixth embodiment.

According to an eighth embodiment of the invention, there is provided a process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising transforming or transfecting said plant cell with a DNA sequence of the third embodiment or a chimeric gene of the fourth embodiment or a plasmid or vector of the fifth embodiment, and causing said DNA sequence to express said polynucleotide or said polypeptide.

According to a ninth embodiment of the invention, there is provided a process for increasing genetic variation in a plant comprising obtaining a hybrid plant from a first plant and a second plant, or cells thereof, said first and second plants being genetically different; altering the mismatch repair system in said hybrid plant; permitting said hybrid plant to self-fertilise and produce offspring plants; and screening said offspring plants for plants in which homeologous recombination has occurred. For example, homeologous recombination may be evidenced by new genetic linkage of a desired characteristic trait or of a gene which contributes to a desired characteristic trait.

According to a tenth embodiment of the invention there is provided a process for obtaining a plant having a desired characteristic, comprising altering the mismatch repair system in a plant, cell or plurality of cells of a plant which does not have said desired characteristic, permitting mutations to persist in said cells to produce mutated plant cells, deriving plants from said mutated plant cells, and screening said plants for a plant having said desired characteristic.

In a preferred form of the ninth and tenth embodiments of the invention, the step of altering the mismatch repair system comprises introducing into said hybrid plant, plant, cell or cells a chimeric gene of the fourth embodiment and permitting the chimeric gene to express a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence in a mismatch repair gene of the hybrid plant, plant, cell or cells, or a polypeptide capable of disrupting the DNA mismatch repair system of the hybrid plant or cells.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 provides a diagrammatic representation of the primer sequences used to isolate *AtMSH3*.

Figure 2 is a plasmid map of clone 52, showing restriction enzyme cleavage sites in the 5' half of the full-length cDNA for *AtMSH3*.

Figure 3 is a plasmid map of clone 13, showing restriction enzyme cleavage sites in the 3' half of the full-length cDNA for *AtMSH3*.

Figure 4 is a sequence listing of the coding sequence of *AtMSH3*, together with a deduced sequence of the encoded polypeptide.

Figure 5 is a protein alignment of yeast (*Saccharomyces cerevisiae*) and *Arabidopsis thaliana* MSH3 protein.

Figure 6 provides a diagrammatic representation of the primer sequences used to isolate *AtMSH6*.

Figure 7 is a plasmid map of clone 43, showing restriction enzyme cleavage sites in the 5' half of the full-length cDNA for *AtMSH6*.

Figure 8 is a plasmid map of clone 62, showing restriction enzyme cleavage sites in the 3' half of the full-length cDNA for *AtMSH6*.

Figure 9 is a sequence listing of the coding sequence of *AtMSH6*, together with a deduced sequence of the encoded polypeptide.

Figure 10 is a protein alignment of yeast (*Saccharomyces cerevisiae*) and *Arabidopsis thaliana* MSH6 protein.

Figure 11 is a genomic sequence listing of *AtMSH6*.

Figure 12 is a plasmid map of plasmid pPF13.

Figure 13 is a plasmid map of plasmid pPF14.

Figure 14 is a plasmid map of plasmid pCW186.

Figure 15 is a plasmid map of plasmid pCW187.

Figure 16 is a diagrammatic representation of an antisense gene construction for use in homeologous meiotic recombination.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is based on the inventors' discovery that there exist in higher plants genes which are homologous to MMR genes in *E. coli*, and to MMR genes in yeasts and humans.

Thus, the inventors have identified genes, herein designated *AtMSH3* and *AtMSH6*, of the plant *Arabidopsis thaliana* which encode the proteins AtMSH3 and AtMSH6. These plant proteins are homologous to yMSH3 and yMSH6, respectively. The present inventors have isolated cDNAs encoding the proteins AtMSH3 and AtMSH6 and have isolated the complete gene encoding AtMSH6. Given the teaching herein, other genes may be obtained (including genes of other plants) which are involved in DNA mismatch repair in plants, including other genes which encode polypeptides homologous to MMR proteins of yeasts or humans. For example, given the teaching herein, genes of members of the *Brassicaceae* family or of other unrelated families, for example the *Poaceae*, the *Solanaceae*, the *Asteraceae*, the *Malvaceae*, the *Fabaceae* and the *Cucurbitaceae* family, and which encode polypeptides homologous to MMR proteins of yeasts or humans may be obtained.

A procedure which may be followed to obtain genes *AtMSH3* and *AtMSH6* is described in Example 1. Essentially the same technique may be applied to cDNA obtained by reverse transcription of RNA from other plants. Alternatively, given the sequence information disclosed herein, other degenerate oligonucleotide primers may be designed and obtained for use in isolating sequences of plant mismatch repair genes which are homologous to *AtMSH3* or *AtMSH6*, from other plants.

The successful gene isolation disclosed herein demonstrates for the first time the existence of MMR in higher plants and indicates the presence of other plant MMR genes. For example, genes encoding the plant homologs of MSH1, MSH2, MSH4, MSH5, PMS1, PMS2 and MLH1 may be identified given the teaching herein. Such genes, as well as those specifically described herein, separately or in combination, are useful in

manipulating the plant MMR for plant breeding purposes. Thus, for example, the plant MMR may be altered by including in a plant cell a polynucleotide sequence as defined herein above with reference to the third embodiment of the invention, and causing the polynucleotide sequence to express either a polynucleotide which disables a plant MMR gene, or a polypeptide which disrupts the plant's MMR system.

The DNA molecule of the third embodiment of the invention includes a polynucleotide sequence (herein referred to as a MMR altering gene) which may for example encode sense, antisense or ribozyme molecules characterised by sufficient base sequence similarity or complementarity to the gene to be altered to permit the antisense or ribozyme molecule to hybridise with the plant MMR gene in vivo or to permit the sense molecule to participate in co-suppression. Alternatively, the MMR altering gene may encode a protein or proteins which interfere with the activity of a plant MMR protein and thus disrupt the plant's MMR system. For example, such encoded proteins may be antibodies or other proteins capable of interfering with MMR protein function.

A chimeric gene of the fourth embodiment, incorporating a MMR altering gene, may be prepared by methods which are known in the art. Similarly, the MMR altering gene may be introduced into a plant cell, regenerating tissue or whole plant by techniques known in the art as being suitable for plant transformation, or by crossing. Known transformation techniques include *Agrobacterium tumefaciens* or *A. rhizogenes* mediated gene transfer, ballistic and chemical methods, and electroporation of protoplasts.

The MMR altering gene or genes are typically expressed from suitable promoters. Suitable promoters may direct constitutive expression, such as the 35S or the *NOS* promoter. Usually, however, the promoter will direct either inducible or tissue specific (e.g. callus; embryonic tissue; etc.), cell type specific (e.g. protoplasts; meiocytes; etc.) or developmental (e.g. embryo) expression of the altering gene or genes, in order for the MMR system to function in tissue types or cell types, or at developmental stages of the plant, in which it is not desirable for the MMR system to be altered. Using such promoters, therefore, the activity of a MMR altering gene may be limited to a specific stage during plant development or it may be altered by controlling conditions external to the plant, and the deleterious effects of a permanently disabled or altered DNA mismatch repair system in a plant may be avoided. Examples of suitable promoters which are not constitutive are known in the art and include inducible promoters such as *PR*1a (reviewed by Gatz, 1997, Annual Rev. Plant Phys. Plant Mol. Biol. 48: 89), tissue specific promoters such as *AoPR1* (Sabahattin et al., 1993, Biotechnology 11: 218), and cell-type specific promoters such as *DMC*1.

A chimaeric gene in accordance with the invention may further be physically linked to one or more selection markers such as genes which confer phenotypic traits such as herbicide resistance, antibiotic resistance or disease resistance, or which confer some

other recognisable trait such as male sterility, male fertility, grain size, colour, growth rate, flowering time, ripening time, etc.

The process of the tenth embodiment of the invention provides, for example, a process for generating intraspecies genetic variation by altering the mismatch repair system in a plant cell, in regenerating plant tissue or in a whole plant. The plant cell, regenerating tissue or whole plant includes and expresses one or more MMR altering genes which are capable of altering mismatch repair in the plant cell, regenerating tissue or whole plant. Alteration of MMR may be achieved, for example, by inactivating the genes encoding plant MSH3 and/or plant MSH6. It is preferred to inactivate the plant MSH3 and MSH6 encoding genes at the same time and in the same plant cell, regenerating tissue or whole plant. Typically in this preferred form of the invention inactivation of either plant MSH3 or MSH6 alone is insufficient to substantially alter the plant's mismatch repair system and only when both MSH3 and MSH6 are inactivated simultaneously is the plant's mismatch repair system sufficiently altered to prevent the MMR system from recognising base pair mismatches as a result of DNA replication errors, DNA damage, or oligonucleotide induced site-specific mutagenesis. However, in some applications of the invention, inactivation of only one gene may also be used to cause genomic instability or increase the efficiency of site-specific mutagenesis.

If desired, the MMR altering gene or genes may later be rendered non-functional or ineffective, or may be removed from the genome of the plant cell, regenerating tissue or whole plant in order to restore mismatch repair in the plant cell, regenerating tissue or whole plant. The MMR altering gene or genes may be inactivated by means of known gene inactivation tools, such as ribozymes, or may be removed from the genome using gene elimination systems known in the art, such as *CRE/LOX*. It is preferred to render two genes, whose gene products combine to incapacitate MMR, ineffective by separating the altering genes through segregation. Therefore, in a preferred embodiment of the invention a first plant cell or plant is generated in which only plant *MSH3* is incapacitated, and a second plant cell or plant is generated in which only plant *MSH6* is incapacitated. The combination of both genomes, for example by crossing, then produces significant MMR deficiency in those cells or plants which have inherited both altering genes. If the altering genes are expressed from unlinked loci, gene segregation restores MMR activity in the progeny of the cells or plants.

In a process of the ninth embodiment of this invention, homeologous recombination is enhanced between different genomes, chromosomes or genes in plant cells or plants by altering MMR in said plant cells or plants. Such genomes, chromosomes or genes are characterised in that they originate from different plant families, genera, species, subspecies, plant varieties or lines. Hybrid plant cells or hybrid plants may be produced by crossing, by cell fusion or by other techniques known in the art. These plant cells or

plants are further characterised by expressing one or more genes that are capable of altering mismatch repair in the plant cell or plants.

In the process of the ninth embodiment, the homeologous recombination is typically for the purpose of introducing a desired characteristic in the hybrid plant. In this typicl application of the process of the ninth embodiment, and iIn the process of the tenth embodiment the desired characteristic may be any characteristic which is of value to the plant breeder. Examples of such characteristics are well known in the art and include altered composition or quality of leaf or seed derived storage products (e.g. oil, starch, protein), altered composition or quality of cell walls (e.g. decrease in lignin content), altered growth rate, prolonged flowering, increased plant yield or grain yield, altered plant morphology, tolerance to or improved performance under environmental stresses of various kinds, etc.

In a preferred form of the tenth embodiment, an MMR altering gene is co-introduced along with the homeologous genome, chromosome or gene of another plant cell or plant into an MMR proficient plant cell or MMR proficient plant to produce a hybrid plant cell or hybrid plant in which homeologous recombination can occur. Suitably, the MMR proficient plant cell or MMR proficient plant may also include an MMR altering gene. For example a gene capable of inactivating plant MSH3 may be co-introduced along with the homeologous genome, chromosome or gene of another plant cell or plant into an MMR proficient plant cell or MMR proficient plant in which MSH6 is inactivated. A resultant hybrid plant in which homeologous recombination occurs will include both the MSH3 and MSH6 altering genes and its MMR system will therefore be inactivated.

In this form of the invention, if hybrid plants are to be produced by crossing, the MMR altering gene preferably originates from the male parent, thus ensuring that the MMR altering gene is always introduced and is not present in the recipient cell. That is, the MMR of the recipient cell, prior to introduction of the MMR altering gene, is typically proficient. Alternatively, if an MMR altering gene is present in a recipient cell it may be ineffective or inefficient on its own, or it may be linked to an inducible or tissue specific or cell type specific promoter which only renders the MMR altering gene active under limited conditions.

Thus, in a preferred form of the process of the ninth embodiment, the MMR system of the hybrid plant is initially unaltered. In this form of the process, the step of altering the mismatch repair system may comprise introducing into the hybrid plant, or cells thereof, a MMR altering gene, such as by *Agrobacterium tumefaciens* or *A. rhizogenes* mediated gene transfer, ballistic and chemical methods, and electroporation of protoplasts.

The MMR altering gene or genes are typically expressed from suitable promoters, as described above. Preferably, the promoter is transcriptionally active in mitotically and

meiotically active tissue and/or cells to ensure MMR alteration after chromosome pairing at mitosis and meiosis, respectively. The preferred timing for MMR alteration is at meiosis, because recombinant genomes, chromosomes or genes are directly transmitted to the progeny. A suitable meiocyte specific promoter is for example the *DMC*1 promoter from *Arabidopsis thaliana* ssp. *L.er.* (Klimyuk and Jones, 1997, Plant J. 11, 1-14). However, mitotic homeologous recombination is also a desirable outcome as somatic recombination events can be transmitted to offspring due to the totipotency of plant cells and the lack of predetermined germ cells in plants.

If desired, the MMR altering gene or genes may later be rendered non-functional or ineffective, or may be removed from the hybrid plant or hybrid plant cells, in order to restore mismatch repair in the hybrid plant or hybrid plant cells. The MMR altering gene or genes may be inactivated by means of known gene inactivation tools as described herein above.

EXAMPLES

Example 1. Cloning of the *AtMSH3* and *AtMSH6* coding sequences

Isolation of partial *AtMSH3* and *AtMSH6* consensus sequences

Degenerate oligonucleotides UPMU and DOMU

UPMU CTGGATCCACIGGICCIAA(C/T)ATG

DOMU CTGGATCC(A/G)TA(A/G)TGIGTI(A/G)C(A/G)AA

were used to isolate *AtMSH3* and *AtMSH6* sequences by PCR amplification.

Primers UPMU and DOMU correspond to conserved amino acid sequences of the proteins MutS (*E. coli* and *S. typhimurium*), HexA (*S. pneumoniae*), Rep1 (mouse) and Duc1 (human). The conserved regions to which they are targeted are TGPNM for UPMU (amino acid positions 852-856 for AtMSH6 and 817-821 for AtMSH3) and FATHY for DOMU (amino acid positions 965-969 for AtMSH6 and 929-933 for AtMSH3.) These primers have been used to isolate MSH2 and MSH1 from yeast (Reenan and Kolodner, Genetics 132: 963-973 (1992)) and MSH2 from *Xenopus* and mouse (Varlet et al., Nuc. Acids Res. 22:5723-5728 (1994)).

Template single strand cDNA was produced by reverse transcription of 2 μg total RNA from a cell suspension culture of *Arabidopsis thaliana* ecotype Columbia (Axelos et al. 1989, Mol. Gen. Genetics 219: 106-112). The PCR reaction was performed under the following conditions in a final volume of 100μl: 0.2mM dNTP, 1μM each primer, 1XPCR buffer, 1u *Taq* DNA polymerase (Appligene) in the presence of template cDNA. PCR parameters were 5 minutes at 94°C, followed by 30 cycles of 40 seconds at 95°C, 90 seconds at 45°C, 1 minute at 72°C. The amplification products were cloned into pGEM-T vector (Promega) and sequenced. Two different clones were isolated, S5 (350bp) was homologous to *MSH3*, S8 (327bp) was homologous to *MSH6*. Complete cDNA sequences were then isolated according to the Marathon cDNA amplification kit

procedure (Clontech). In summary, this procedure involves producing double stranded cDNA by reverse transcription of 2μg polyA+ RNA from the cell suspension culture of *Arabidopsis*. Adaptors are ligated on each side of the cDNA. The ligated cDNA is used as a template for 5' and 3' RACE PCR reactions in the presence of primers that are specific for the adaptor on one side (AP1 and AP2), and specific for the targeted gene on the other side. A 5' and a 3' fragment that overlap are thus produced for each gene. The complete gene coding sequence can be reconstituted taking advantage of a unique restriction site, if available, in the overlapping region. Specific details of this procedure as it was used to isolate *AtMSH3* and *AtMSH6* coding regions, are as follows.

Isolation of *AtMSH3* complete coding sequence

From the sequence of clone S5, primer 636 was designed:

636 TGCTAGTGCCTCTTGCAAGCTCAT.

Primer AP1 is complementary to a portion of an adaptor sequence which had been ligated to the 5' and 3' ends of *Arabidopsis* cDNA:

AP1 CCATCCTAATACGACTCACTATAGGGC.

PCR performed on the ligated cDNA with primers 636 and AP1 for the 5' RACE PCR was followed by a second round of amplification with the nested primers AP2 and S525

AP2 ACTCACTATAGGGCTCGAGCGGC

S525 AGGTTCTGATTATGTGTGACGCTTTACTTA

(the latter was also designed to correspond to a part of the sequence of clone S5) and produced a 2720bp DNA fragment. Figure 1 provides a diagrammatic representation of the primer sequences used to isolate *AtMSH3*. Another primer (S51)

S51 GGATCGGGTACTGGGTTTTGAGTGTGAGG

was designed closer to the 5' border and permitted the determination of 99bp upstream to the ATG initiation codon. For the 3' RACE PCR, a first PCR reaction was performed with primers AP1 and 635,

635 GCACGTGCTTGATGGTGTTTTCAC

followed by a second round of amplification, using the nested primers AP2 and S523

S523 TCAGACAGTATCCAGCATGGCAGAAGTA

which produced a DNA fragment of 890bp. Both DNA fragments were subcloned into pGEM-T and sequenced. Since PCR amplification using the Expand Long Template PCR System (Boehringer-Mannheim) produced errors in the sequence, new oligonucleotides were designed to isolate those sequences again by PCR, but with the high fidelity DNA polymerase *Pfu*. PCR with primers 1S5 and S53

1S5 ATCCCGGGATGGGCAAGCAAAAGCAGCAGACGA

S53 GACAAAGAGCGAAATGAGGCCCCTTGG

amplified the 1244bp fragment clone 52 (cloned into pUC18/SmaI). PCR with primers S52 and 2S5

2S5 ATCCCGGGTCAAAATGAACAAGTTGGTTTTAGTC

S52 GCCACATCTGACTGTTCAAGCCCTCGC

amplified the 2104bp clone 13 (cloned into pUC18/SmaI). These two clones were ligated after digestion by the BamH1 restriction enzyme for which a unique site is present in their overlapping region. The remaining primers referred to in Figure 1 are as follows:

S51 GGATCGGGTACTGGGTTTTGAGTGTGAGG

S525 AGGTTCTGATTATGTGTGACGCTTTACTTA

Figures 2 and 3 provide plasmid maps of clones 52 and 13 respectively, showing restriction enzyme cleavage sites. The complete *AtMSH3* coding sequence is 3246bp long and is shown in Figure 4 together with the deduced sequence of the encoded polypeptide. *AtMSH3* is clearly homologous to the yeast and mouse *MSH3* genes. A sequence alignment of polypeptides encoded by *AtMSH3* and that encoded by *Saccharomyces cerevisiae MSH3* is set out in Figure 5.

Isolation of the *AtMSH6* complete coding sequence and genomic sequences

The same procedure allowed isolation of the *AtMSH6* cDNA. Figure 6 provides a diagrammatic representation of the primer sequences used to isolate *AtMSH6*. For the 5' RACE PCR, primers 638 and AP1

638 TCTCTACCAGGTGACGAAAAACCG

allowed the amplification of a 2889 DNA fragment. Primer S81

S81 CGTCGCCTTTAGCATCCCCTTCCTTCAC

helped define the 142bp upstream to the ATG initiation codon. On the 3' side, RACE PCR was initially performed with primers S823 and AP1,

S823 GCTTGGCGCATCTAATAGAATCATGACAGG

and then with the nested primers 637 and AP2,

637 GACAGCGTCAGTTCTTCAGAATGC

to produce a 774bp DNA fragment. As for *AtMSH3*, those fragments were cloned and sequenced. Re-isolation of the DNA sequence using the high fidelity *Pfu* polymerase and newly designed primers 1S8 and S83 (for the 5' side) led to a 2182 bp DNA fragment identified as clone 43 (cloned in pUC18/SmaI), and a 1379bp clone identified as clone 62 (also cloned in pUC18/SmaI).

1S8 ATCCCGGGATGCAGCGCCAGAGATCGATTTTGT

2S8 ATCCCGGGTTATTTGGGAACACAGTAAGAGGATT

S82 GCGTTCGATCATCAGCCTCTGTGTTGC

S83 CGCTATCTATGGCTGCTTCGAATTGAG

Figures 7 and 8 provide plasmid maps of clones 43 and 62 respectively, showing restriction enzyme cleavage sites. Clones 43 and 62 were digested by the *Xmn*1 restriction enzyme for which a unique site is present in their overlapping region and then ligated. The complete *AtMSH6* coding sequence is 3330bp long and is shown in Figure 9 together with the deduced sequence of the encoded polypeptide. *AtMSH6* is clearly

homologous to the yeast and mouse *MSH6*genes. A sequence alignment of polypeptides encoded by *AtMSH6* and that encoded by *Saccharomyces cerevisiae MSH6* is set out in Figure 10.

An *AtMSH6* genomic sequence was also isolated from a genomic DNA library constituted after partial *Sau3*AI digestion of DNA from the *Arabidopsis* cell suspension. 8062bp were sequenced that covered the *AtMSH6* gene and show colinearity with the cDNA. 16 introns are found scattered along the gene. The complete genomic sequence is shown in Figure 11.

Example 2. A measure of somatic variation in MMR deficient plants

Constructs

Constructs with antisense *AtMSH3* or antisense *AtMSH6* or both *AtMSH3/AtMSH6* under the control of a single 35S promoter have been inserted into the binary vector pPZP121 between the right and left borders of the T-DNA. The pPZP121 plasmid confers chloramphenicol resistance to *Escherichia coli* or *Agrobacterium tumefaciens* bacteria. The *aacC1* gene is carried by the T-DNA and allows selection of transformed plant cells on gentamycine (Hajdukiewicz et al., Plant Mol. Biol. 25, 989-994). For the purpose of expressing antisense constructs, a 35S promoter/terminator cassette with a polylinker was introduced into pPZP121. The 3' ends of the considered genes have been chosen since this region seems more efficient for antisense inhibition. For *AtMSH3* this corresponds to clone 13 (2104bp), for *AtMSH6* this corresponds to clone 62 (1379bp). Clone 13 comprises 2104bp of the 3' region that were cut off the pUC18 vector by Sal1/Sst1 restriction, blunted with T4 DNA polymerase and ligated into the T4 DNA polymerase blunted BamH1 site of pPZP121/35S, creating clone pPF13. The same procedure was followed for the 3' region of AtMH6 clone 62 (1379bp) thus creating plasmid pPF14. For the double constructs, the 3' region (from clone 62) of AtMSH6 was introduced ahead of the AtMSH3 region into pPF13 creating pCW186 and reciprocally, the 3' region of *AtMSH3* (from clone 13) was introduced ahead of *AtMSH6* into pPF14, creating pCW187. Figures 12-15 provide plasmid maps of plasmids pPF13, pPF14, pCW 186 and pCW187, respectively.

Plant cell transformation

The constructs are introduced into *Agrobacterium* by electroporation. Plant cells are then transformed by co-cultivation. A suspension culture of *Arabidopsis thaliana* cells that has been established by Axelos et al. (1992, Plant Physiol. Biochem. 30, 1-6) may be used. One day old freshly subcultured cells are diluted five times into AT medium (Gamborg B5 medium, 30g/l sucrose, 200µg/l NAA). 10µl of saturated *Agrobacterium* containing the transforming T-DNA constructs are added to 10ml diluted cells in a 100ml erlenmeyer. The co-cultivation is agitated slowly (80rpm) for 2 days. The cells are then washed 3 times into AT medium and finally resuspended in the same

initial volume (10ml). The culture is agitated for 3 days to allow expression before plating on selection plates (AT/BactoAgar 8g/l+gentamycine 50μg/ml). Transformed individual calli are isolated 3 weeks later.

Microsatellite analysis

Microsatellites have been described in *Arabidopsis* (Bell and Ecker, 1994, Genomics 19, 137-144). We chose to study instability of microsatellites that are polymorphic amongst different ecotypes. DNA is extracted from the transformed calli. Specific primers have been defined that are used to amplify the microsatellite sequence. One of the two primers is previously P^{32} labelled by T4 kinase. In case of a polymorphic variation, new PCR products appear that do not follow the expected pattern of migration on a polyacrylamide gel. This is a commonly observed feature for MMR deficient cells in yeast or mammalian cells.

Forward mutagenesis assay

A single copy of the *codA* gene has been introduced in the cells prior to transformation by the antisense constructs. This *codA* gene renders the cells sensitive to the base analog 5-fluoro-cytosine (Perera et al., 1993, Plant Mol. Biol. 23, 793-799). Inactivation of MMR is expected to increase mutagenesis. Newly arising mutations will eventually inactivate the codA gene conferring resistance to 5 fluoro-cytosine. The *codA* gene from resistant clones will be isolated by PCR and sequenced to check for the occurrence of mutations in its coding sequence.

Example 3. Homeologous meiotic recombination in *Arabidopsis thaliana*

A. Construction of a meiocyte specific gene expression cassette comprising the *DMC1* promoter and the *NOS* terminator

A 1.8 kb DNA fragment comprising the 3' terminal part of the meiocyte specific *DMC1* promoter (Klimyuk and Jones, 1997, Plant J. 11, 1-14) was isolated by PCR from purified genomic DNA of *Arabidopsis thaliana* (ssp. *Landsber erecta* "Ler"). The forward PCR primer (DMC1a) contained the sequence

acgcgtcgacGAATTCGCAAGTGGGG

and introduced a *SalI* cloning site at the 5' end of the promoter fragment. The reverse PCR primer (DMC1b) contained the sequence

tccatggagatctcccgggtacCGATTTGCTTCGAGGG

introducing a polylinker region at the 3' end of the promoter fragment. The PCR reaction was carried out with VENT DNA Polymerase (NEB) over 25 cycles using the following cycling protocol: 1 minute at 94°C, 1 minute at 54°C, 2 minutes at 72°C.

The PCR reaction yielded a blunt ended DNA fragment which was digested with restriction enzyme *SalI* and was cloned into the cleavage sites of restriction enzymes *SalI* and *SmaI* in plasmid p2030, a pUC118 derivative containing the *SacI-EcoRI* NOS

terminator fragment of pBIN19. The cloning yielded plasmid p2031, containing the *DMC*1 promoter-polylinker-NOS terminator expression cassette depicted in Figure 16.

B. Construction of an *MSH*3 antisense gene under the control of the *DMC*1 promoter

A 2.1 kb DNA fragment encoding the carboxyterminal part of AtMSH3 was removed from the polylinker of clone 13 described in Example 1 after (i) digestion with *Kpn*I, (ii) blunting of the DNA ends generated by *Kpn*I and (iii) digestion with *Bam*HI. The isolated fragment was then cloned in antisense orientation downstream of the *DMC*1 promoter in plasmid p2031, which had been digested with restriction enzymes *Sma*I and *Bgl*II. This cloning yielded plasmid p2033 (Figure 16).

After digestion of p2033 with *Eco*RI, a 4.1 kb DNA fragment was recovered comprising the *DMC*1 promoter, the partial *MSH*3 cDNA in antisense orientation with respect to the promoter and the *NOS* terminator. This fragment was cloned into the *Eco*RI restriction site of plant transformation vector pNOS-Hyg-SCV to yield plasmid p3242 (Figure 16).

C. Construction of a combined *MSH3/MSH6* antisense gene under the control of a single *DMC*1 promoter

A 1 kb DNA fragment encoding the carboxyterminal part of ATMSH6 is isolated from clone 62 described in Example 1 after digestion of clone 62 plasmid DNA with *Bam*HI, which cleaves in the 5' polylinker region flanking the partial cDNA, and with *Eco*RI, which cleaves within the cDNA. The isolated fragment is treated with *Klenow* enzyme to blunt both its ends and is cloned into the recipient plasmid p2033 or p3242. For the purpose of cloning, the recipient plasmid may be cleaved with either *Ava*I or *Nco*I and can be blunted with *Klenow* enzyme to produce blunt acceptor ends for fragment cloning. This cloning yields two opposite orientations of cloned fragment DNA with respect to the *DMC*1 promoter. These can be identified by diagnostic digestion with restriction enzymes *Sca*I or *Xmn*I in conjunction with *Sac*I. The selected construct contains the *DMC*1 promoter, the combined partial cDNAs for *AtMSH3* and *AtMSH6* (both cloned in antisense orientation with respect to the *DMC*1 promoter) and the *NOS* terminator. If the recipient plasmid is p2033, the combined antisense gene under control the single *DMC*1 promoter is recovered from the vector after *Eco*RI digestion and is cloned into the *Eco*RI restriction site of pNOS-Hyg-SCV.

D. Selection of Recombination markers on homeologous chromosomes of *Arabidopsis thaliana* subspecies *Landsberg erecta* (Ler), *Columbia* (Col) and *C24*, respectively

D(i). Visual recombination markers in *Arabidopsis th.* subspecies *C24*:

Agrobacterium mediated transformation with a T-DNA containing a *35S-GUS* gene, inactivated by insertion of a *35S-Ac* transposable element (Finnegan et al., 1993, Plant Mol. Biol. 22, 625-633), had yielded a *C24* line in which the T-DNA construct was

integrated into chromosome 2. Genetic and molecular analysis of this line shows that the *Ac* transposon had excised from its T-DNA locus thereby restoring *GUS* activity, but had re-inserted into the chromosome at a distance of 16.4 cM, where it stayed fixed (due to disablement of *Ac*) within the *chlorina* gene. Insertional inactivation of the *chlorina* gene caused a bleached phenotype in those plants that were homozygous for this mutation. Because of the two linked phenotypic markers, *chlorina3A::Ac* and *GUS*, this *G24* line was used in crosses to wildtype Ler for analysis of meiotic homeologous recombination on chromosome 2 in conjunction with molecular recombination markers.

D(ii). Visual recombination markers in *Arabidopsis th. Ler*:

The Ler line NW1 (obtained from NASC, Nottingham, UK) contains one recessive visual marker per chromosome, i.e. *an-1* on Chr.1, *py-1* on Chr.2, *gl1-1* on Chr.3, *cer2-1* on Chr.4, and *ms1-1* on Chr.5. This line was used in crosses to wildtype *C24* for analysis of meiotic homeologous recombination on chromosomes 1-5 in conjunction with molecular recombination markers.

D(iii) Molecular recombination markers in *Col*, *Ler* and *C24*:

The genome of *Arabidopsis thaliana* is interspersed with unique base sequences arranged as simple tandem repeats. Allelic repeats can vary in length between different *Arabidopsis* subspecies and when amplified by PCR yield diagnostic DNA products of different length named Simple Sequence Length Polymorphisms (SSLPs). Many SSLPs have been genetically mapped and have been assigned to unique chromosome locations on the recombinant inbred map (Bell and Ecker, 1994, Genomics 19, 137-144; Lister and Deans lines, Weeds World 4i, May 1997).

In Table 1 are listed 28 mapped and established SSLPs between *Ler* and *Col*. A number of PCR primer pairs are described herein which also yielded SSLPs between *C24* and *Ler* (19 SSLPs) or between *C24* and *Col* (25 SSLPs), respectively. Polymorphic SSLPs can be used as molecular markers in the analysis of homeologous recombination between genomes from these subspecies.

The PCR reactions (25 μL) were carried out over 35 cycles (15 seconds at 94°C, 30 seconds at 55°C and 30 seconds at 72°C), with 0.25 U Taq DNA polymerase and 0.6 μg genomic DNA in reaction buffer containing 2 mM $MgCl_2$. PCR products were separated by agarose gel electrophoresis (4% ultra high resolution agarose) and visualised by ethidiumbromide staining. The results from the PCR experiments are summarised in Table 1, which also shows the sequence of PCR primers, primer annealing temperature, and chromosome location of SSLP (with respect to the RI map of May 1997, Weeds World 4i).

E. Production of hybrid plants

C24 plants heterozygous for *chlorina3A:Ac/GUS* were crossed as male to emasculated wildtype *Ler* to produce *Ler / C24(chlorina3A, GUS)* hybrid seeds.

Ler plants homozygous for the five chromosome markers were male sterile (*ms1-1*) and were crossed without emasculation to wildtype *C24* to produce *Ler(an-1, py-1, gl1-1, cer2-1, ms1-1)/C24* hybrid seeds.

F. Introduction of *MSH3* and *MSH3/6* antisense genes into hybrid plants and analysis of meiotic homeologous recombination

The plant transformation vectors comprising the antisense genes described in (B) and (C) above are introduced into *Agrobacterium tumefaciens* strain *AGL*1 (Lazo et al., 1991, Bio/Technology 9, 963-967) by electroporation. Recombinant *Agrobacterium* clones are selected on LB medium containing 50 mg/L rifampicin and 50 mg/L kanamycin. Selected clones are used to infect roots of *Arabidopsis* hybrid plants (described in (E) above) using the root transformation protocol of Valvekens et al. (1988, PNAS 85, 5536-5540) except that shoot and root inducing media contain hygromycin (10 mg/L) instead of kanamycin.

Plants regenerated from roots of hybrid plants are genetic clones of root donating plants and therefore are again genetic hybrids of two *Arabidopsis* subspecies described in (E). However, in contrast to the root donating plants, the regenerated hybrid plants also contain the introduced transgene and the co-introduced hygromycin resistance gene with the latter allowing these plants to grow on hygromycin containing culture medium.

Hygromycin resistant plants are then allowed to enter the reproductive phase and to produce gametes by meiotic divisions of microspore and megaspore mothercells. At meiosis, the *DMC*1 promoter is activated and can direct the expression of antisense genes described in (B) and (C) above, leading to decreased *MSH6* and/or *MSH3* gene expression. This in turn depletes the gamete mothercells of MSH6 and/or MSH3 protein, thus causing alteration of MMR during meiotic divisions and increasing the recombination frequency between homeologous chromosomes.

Transgenic plants are then allowed to self-fertilise and to produce seeds. These seeds (F2 seeds with respect to hybrid production), and the plants derived therefrom, carry the homeologous recombination events which can be identified by using the visual and molecular recombination markers described in (D) above.

In case of homeologous recombination between chromosomes of *Ler* and *C24(chlorina3A:Ac, GUS)*, the analysis concentrates on chromosome 2 by selecting plants showing the visual phenotypic marker *chlorina*. This marker thus serves as a reference point as it indicates that respective chromosomes 2 originate from *C24*. Other markers, such as *GUS* or molecular markers, on chromosome 2 may then be used to identify

chromosomal regions which are derived from the *Ler* chromosome as a result of homeologous recombination. F2 plants of control transformants not expressing the antisense gene(s) can be analysed in parallel and the results can be used for comparison to homeologous recombination results obtained in antisense plants.

In case of homeologous recombination between chromosomes of *C24* and *Ler* (*an-1, py-1, gl1-1, cer2-1, ms1-1*), the analysis is similar to that described above, except that the presence of a visual marker on each chromosome facilitates the study of homeologous recombination on all five chromosomes.

TABLE 1: SSLP Markers in *Arabidopsis thaliana* Subspecies

Chromosome	RI Map Position	PCR Primer Pair	Primer Sequence	Tm [°C]	length/COL	length/LER	length/C24
I	2.3	ATEAT1 F	GCCACTGCGTGAATGATATG	57.8	172	162	162
		ATEAT1 R	CGAACAGCCAACATTAATTCCC	58.2			
I	9.3	NGA63 F	AACCAAGGCACAGAAGCG	57.3	111	89	120
		NGA63 R	ACCCAAGTGATCGCCACC	59.6			
I	40.1	NGA248 F	TACCGAACCAAAACACAAAGG	56.1	143	129	no amplific.
		NGA248 R	TCTGTATCTCGGTGAATTCTCC	58.2			
I	81.2	NGA128 F	GGTCTGTTGATGTCGTAAGTCG	60.1	180	190	no amplific.
		NGA128 R	ATCTTGAAACCTTTAGGGAGGG	58.2			
I	81.2	NGA280 F	CTGATCTCACGGACAATAGTGC	60.1	105	85	85
		NGA280 R	GGCTCCATAAAAAGTGCACC	57.8			
I	111.4	NGA111 F	CTCCAGTTGGAAGCTAAAGGG	60	128	162	170
		NGA111 R	TGTTTTTTAGGACAAATGGCG	70			
II	ca. 7.5	NGA168 F	CCTTCACATCCAAAACCCAC	57.8	213	217	208
		NGA168 R	GCACATACCCACAACCAGAA	57.8			

II	ca. 48	NGA1126L	CGCTACGCTTTTCGGTAAAG	57.8	191	199	196
		NGA1126R	GCACAGTCCAAGTCACAACC	59.9			
II	62.2	NGA361L	AAAGAGATGAGAATTTGGAC	51.7	114	120	114
		NGA361R	ACATATCAATATATTAAAGTAGC	49.5			
II	73	NGA168 F	TCGTCTACTGCACTGCCG	59.6	151	135	135
		NGA168 R	GAGGACATGTATAGGAGCCTCG	61.9			
II	ca. 77	AthBIO2 L	TGACCTCCTCTTCCATGGAG	59.9	141	209	139
		AthBIO2 R	TTAACAGAAACCCAAAGCTTTC	54.5			
II	ca. 83	AthUBIQUE L	AGGCAAATGTCCATTTCATTG	54.1	146	148	148
		AthUBIQUE R	ACGACATGGCAGATTTCTCC	57.8			
III	3.4	NGA172 F	AGCTGCTTCCTTATAGCGTCC	60	162	136	140
		NGA172 R	CATCCGAATGCCATTGTTC	55.4			
III	12.8	NGA126 F	GAAAAAACGCTACTTTCGTGG	56.1	119	147	no amplific.
		NGA126 R	CAAGAGCAATATCAAGAGCAGC	58.2			
III	17.5	NGA162 F	CATGCAATTTGCATCTGAGG	55.8	107	89	no amplific.
		NGA162 R	CTCTGTCACTCTTTTCCTCTGG	60.1			

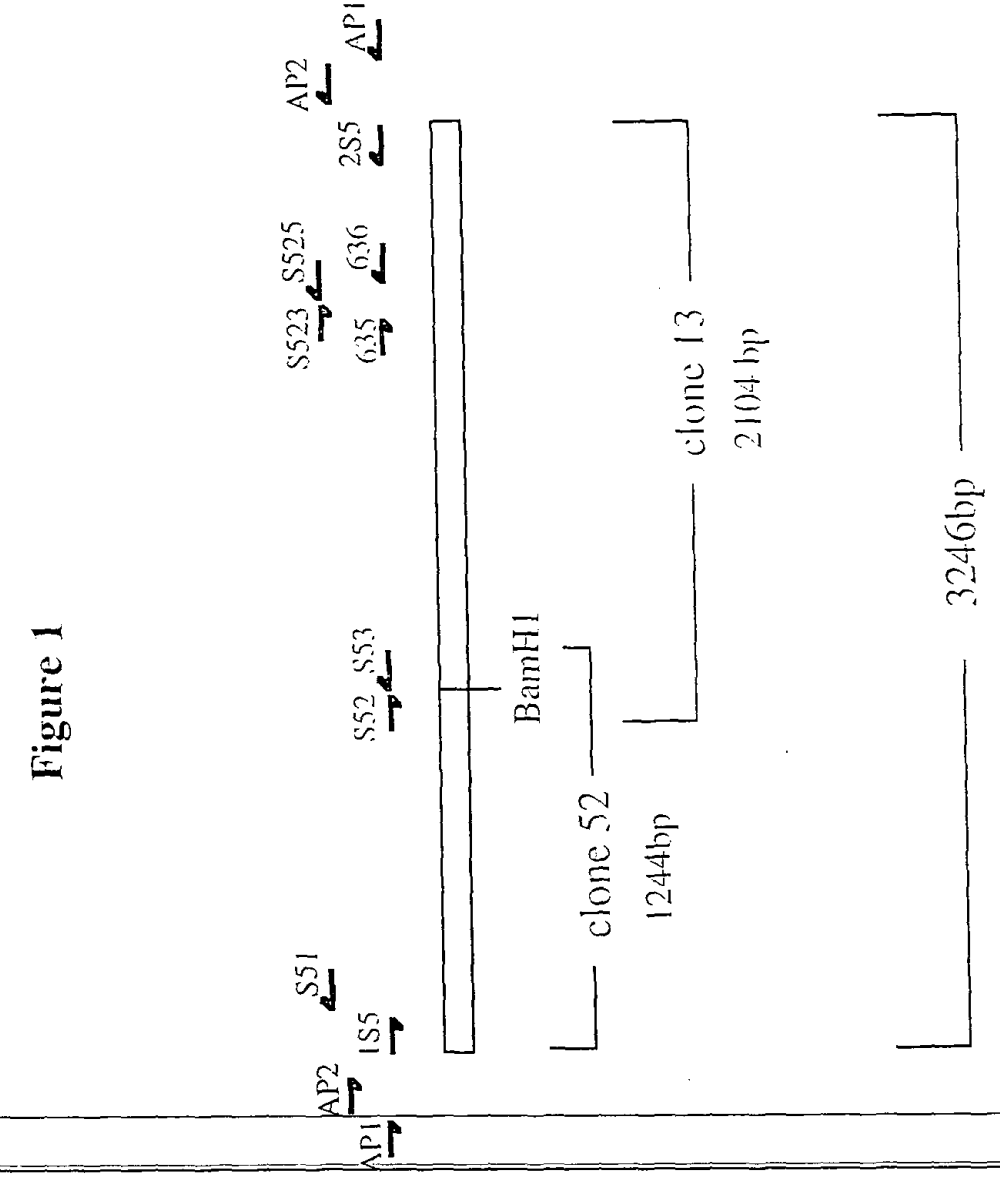
III	81.8	NGA6 F	TGGATTTCTTCCTCTCTTCAC	56.1	143	123	143
		NGA6 R	ATGGAGAAGCTTACACTGATC	56.1			
IV	19.8	NGA12 F	AATGTTGTCCTCCCCTCCTC	59.9	247	234	220
		NGA12 R	TGATGCTCTCTGAAACAAGAGC	58.2			
IV	24.1	NGA8 F	GAGGGCAAATCTTTATTTCGG	56.1	154	198	190
		NGA8 R	TGGCTTTCGTTTATAAACATCC	54.5			
IV	102	NGA1107 L	GCGAAAAAACAAAAAAATCCA	50.2	150	140	140
		NGA1107 R	CGACGAATCGACAGAATTAGG	58			
V	11.8	NGA225 F	GAAATCCAAATCCCAGAGAGG	58	119	189	119
		NGA225 R	TCTCCCCACTAGTTTTGTGTCC	60.1			
V	16.7	NGA249 F	TACCGTCAATTTCATCGCC	55.4	125	115	115
		NGA249 R	GGATCCCTAACTGTAAAATCCC	58.2			
V	19.9	CA72 F	AATCCCAGTAACCAAACACACA	56.3	124	110	110
		CA72 R	CCCAGTCTAACCACGACCAC	61.9			
V	20	NGA151 F	GTTTTGGGAAGTTTTGCTGG	55.8	150	120	130
		NGA151 R	CAGTCTAAAAGCGAGAGTATGATG	58.6			

V	24		NGA106 F	GTTATGGAGTTTCTAGGGCACG	60.1	157	123	130
			NGA106 R	TGCCCCATTTTGTTCTTCTC	55.8			
V	37.8		NGA139 F	AGAGCTACCAGATCCGATGG	59.9	174	132	132
			NGA139 R	GGTTTCGTTTCACTATCCAGG	55.8			
V	50		NGA76 F	GGAGAAAATGTCACTCTCCACC	60.1	231	>250	300
			NGA76 R	AGGCATGGGAGACATTTACG	57.8			
V	61.1		ATHSO191 L	CTCCACCAATCATGCAAATG	55.8	148	156	146
			ATHSO191 R	TGATGTTGATGGAGATGGTCA	53.7			
V	81.7		NGA129 F	TCAGGAGGAACTAAAGTGAGGG	60.1	177	179	172
			NGA129 R	CACACTGAAGATGGTCTTGAGG	60.1			

DATED this TENTH day of OCTOBER 1997
Rhone-Poulenc Agrochimie

Patent Attorneys for the Applicant
SPRUSON & FERGUSON

Figure 1



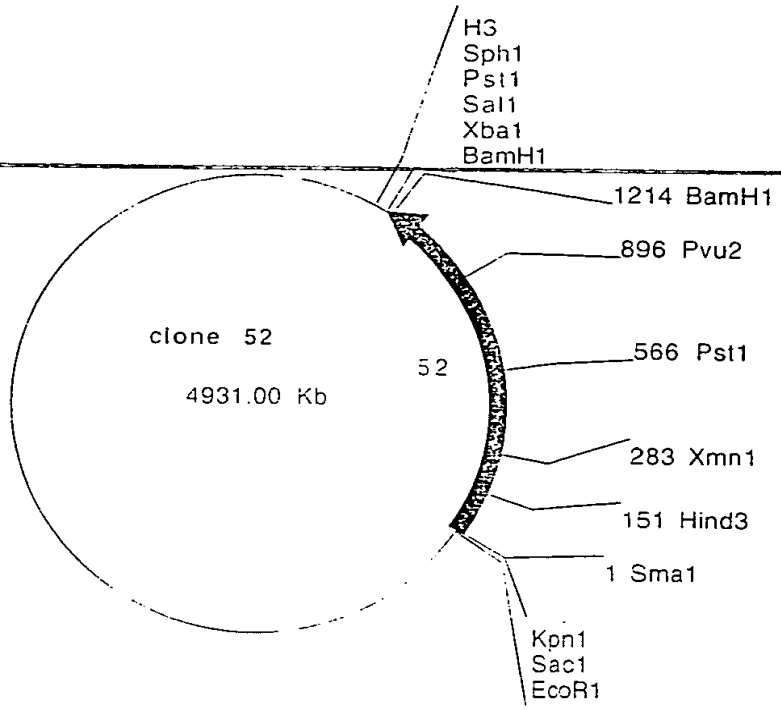


Figure 2

Comments/References: 52= 3' side of S5 (AtMSH3) 1244bp in pUC18/Sma1

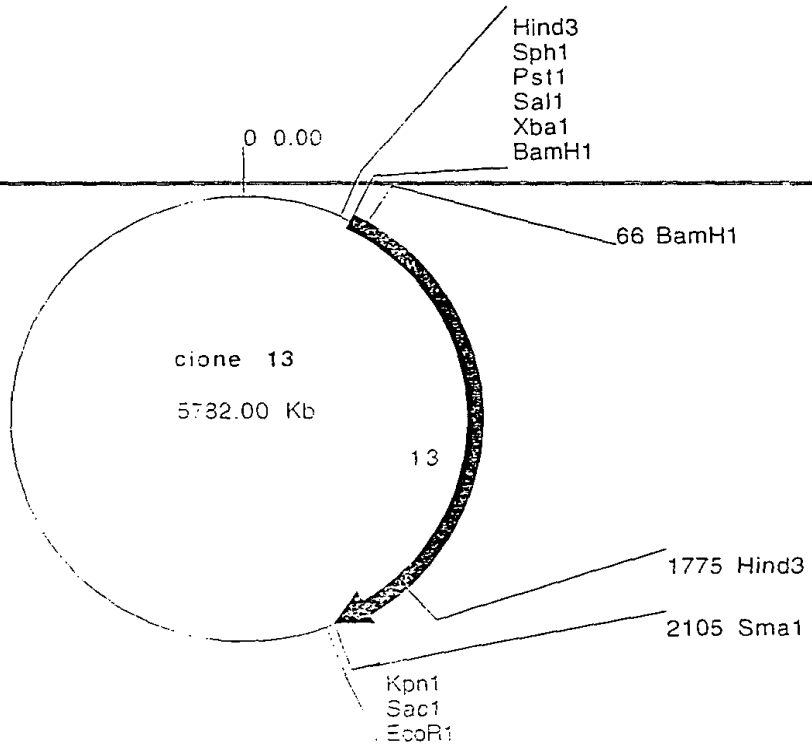


Figure 3

Comments/References: 13 = 3' side of S5 (AtMSH3) 2104bp in pUC18/Sma1

Figure 4

```
   1 cCTAAGAAAGCGCGCGAAAATTGGCAACCCAAGTTCGCCATAGCCACGACCACGACCTTCCATTTCTCTTAAACGGAGGA  80

  81 GATTACGAATAAAGCAATT ATG GGC AAG CAA AAG CAG CAG ACG ATT TCT CGT TTC TTC GCT CCC  144
   1                   M   G   K   Q   K   Q   Q   T   I   S   R   F   F   A   P    15

 145 AAA CCC AAA TCC CCG ACT CAC GAA CCG AAT CCG GTA GCC GAA TCA TCA ACA CCG CCA CCG  204
  16 K   P   K   S   P   T   H   E   P   N   P   V   A   E   S   S   T   P   P   P   35

 205 AAG ATA TCC GCC ACT GTA TCC TTC TCT CCT TCC AAG CGT AAG CTT CTC TCC GAC CAC CTC  264
  36 K   I   S   A   T   V   S   F   S   P   S   K   R   K   L   L   S   D   H   L   55

 265 GCC GCC GCG TCA CCC AAA AAG CCT AAA CTT TCT CCT CAC ACT CAA AAC CCA GTA CCC GAT  324
  56 A   A   A   S   P   K   K   P   K   L   S   P   H   T   Q   N   P   V   P   D   75

 325 CCC AAT TTA CAC CAA AGA TTT CTC CAG AGA TTT CTG GAA CCC TCG CCG GAG GAA TAT GTT  384
  76 P   N   L   H   Q   R   F   L   Q   R   F   L   E   P   S   P   E   E   Y   V   95

 385 CCC GAA ACG TCA TCA TCG AGG AAA TAC ACA CCA TTG GAA CAG CAA GTG GTG GAG CTA AAG  444
  96 P   E   T   S   S   S   R   K   Y   T   P   L   E   Q   Q   V   V   E   L   K  115

 445 AGC AAG TAC CCA GAT GTG GTT TTG ATG GTG GAA GTT GGT TAC AGG TAC AGA TTC TTC GGA  504
 116 S   K   Y   P   D   V   V   L   M   V   E   V   G   Y   R   Y   R   F   F   G  135

 505 GAA GAC GCG GAG ATC GCA GCA CGC GTG TTG GGT ATT TAC GCT CAT ATG GAT CAC AAT TTC  564
 136 E   D   A   E   I   A   A   R   V   L   G   I   Y   A   H   M   D   H   N   F  155

 565 ATG ACG GCG AGT GTG CCA ACA TTT CGA TTG AAT TTC CAT GTG AGA AGA CTG GTG AAT GCA  624
 156 M   T   A   S   V   P   T   F   R   L   N   F   H   V   R   R   L   V   N   A  175

 625 GGA TAC AAG ATT GGT GTA GTG AAG CAG ACT GAA ACT GCA GCC ATT AAG TCC CAT GGT GCA  684
 176 G   Y   K   I   G   V   V   K   Q   T   E   T   A   A   I   K   S   H   G   A  195

 685 AAC CGG ACC GGC CCT TTT TTC CGG GGA CTG TCG GCG TTG TAT ACC AAA GCC ACG CTT GAA  744
 196 N   R   T   G   P   F   F   R   G   L   S   A   L   Y   T   K   A   T   L   E  215

 745 GCG GCT GAG GAT ATA AGT GGT GGT TGT GGT GGT GAA GAA GGT TTT GGT TCA CAG AGT AAT  804
 216 A   A   E   D   I   S   G   G   C   G   G   E   E   G   F   G   S   Q   S   N  235

 805 TTC TTG GTT TGT GTT GTG GAT GAG AGA GTT AAG TCG GAG ACA TTA GGC TGT GGT ATT GAA  864
 236 F   L   V   C   V   V   D   E   R   V   K   S   E   T   L   G   C   G   I   E  255

 865 ATG AGT TTT GAT GTT AGA GTC GGT GTT GTT GGC GTT GAA ATT TCG ACA GGT GAA GTT GTT  924
 256 M   S   F   D   V   R   V   G   V   V   G   V   E   I   S   T   G   E   V   V  275

 925 TAT GAA GAG TTC AAT GAT AAT TTC ATG AGA AGT GGA TTA GAG GCT GTG ATT TCG AGC TTG  984
 276 Y   E   E   F   N   D   N   F   M   R   S   G   L   E   A   V   I   L   S   L  295

 985 TCA CCA GCT GAG CTG TTG CTT GGC CAG CCT CTT TCA CAA CAA ACT GAG AAG TTT TTG GTG 1044
 296 S   P   A   E   L   L   L   G   Q   P   L   S   Q   Q   T   E   K   F   L   V  315

1045 GCA CAT GCT GGA CCT ACC TCA AAC GTT CGA GTG GAA CGT GCC TCA CTG GAT TGT TTC AGC 1104
 316 A   H   A   G   P   T   S   N   V   R   V   E   R   A   S   L   D   C   F   S  335

1105 AAT GGT AAT GCA GTA GAT GAG GTT ATT TCA TTA TGT GAA AAA ATC AGC GCA GGT AAC TTA 1164
 336 N   G   N   A   V   D   E   V   I   S   L   C   E   K   I   S   A   G   N   L  355

1165 GAA GAT GAT AAA GAA ATG AAG CTG GAG GCT GCT GAA AAA GGA ATG TCT TGC TTG ACA GTT 1224
 356 E   D   D   K   E   M   K   L   E   A   A   E   K   G   M   S   C   L   T   V  375

1225 CAT ACA ATT ATG AAC ATG CCA CAT GTG ACT GTT CAA GCC CTC GCC CTA ACG TTT TGC CAT 1284
 376 H   T   I   M   N   M   P   H   L   T   V   Q   A   L   A   L   T   F   C   H  395

1285 CTC AAA CAG TTT GGA TTT GAA AGG ATC CTT TAC CAA GGG GCC TCA TTT CGC TCT TTG TCA 1344
 396 L   K   Q   F   G   F   E   R   I   L   Y   Q   G   A   S   F   R   S   L   S  415

1345 AGT AAC ACA GAG ATG ACT CTC TCA GCC AAT ACT CTG CAA CAG TTG GAG GTT GTG AAA AAT 1404
 416 S   N   T   E   M   T   L   S   A   N   T   L   Q   Q   L   E   V   V   K   N  435

1405 AAT TCA GAT GGA TCG GAA TCT GGC TCC TTA TTC CAT AAT ATG AAT CAC ACA CTT ACA GTA 1464
 436 N   S   D   G   S   E   S   G   S   L   F   H   N   M   N   H   T   L   T   V  455

1465 TAT GGT TCC AGG CTT CTT AGA CAC TGG GTG ACT CAT CCT CTA TGC GAT AGA AAT TTG ATA 1524
 456 Y   G   S   R   L   L   R   H   W   V   T   H   P   L   C   D   R   N   L   I  475

1525 TCT GCT CGG CTT GAT GCT GTT TCT GAG ATT TCT GCT TGC ATG GGA TCT CAT AGT TCT TCC 1584
 476 S   A   R   L   D   A   V   S   E   I   S   A   C   M   G   S   H   S   S   S  495
```

igure 4 (continued)

```
1585 CAG CTC AGC AGT GAG TTG GTT GAA GAA GGT TCT GAG AGA GCA ATT GTA TCA CCT GAG TTT 1644
 496 Q   L   S   S   E   L   V   E   E   G   S   E   R   A   I   V   S   P   E   F   515

1645 TAT CTC GTG CTC TCC TCA GTC TTG ACA GCT ATG TCT AGA TCA TCT GAT ATT CAA CGT GGA 1704
 516 Y   L   V   L   S   S   V   L   T   A   M   S   R   S   S   D   I   Q   R   G   535

1705 ATA ACA AGA ATC TTT CAT CGG ACT GCT AAA GCC ACA GAG TTC ATT GCA GTT ATG GAA GCT 1764
 536 I   T   R   I   F   H   R   T   A   K   A   T   E   F   I   A   V   M   E   A   555

1765 ATT TTA CTT GCG GGG AAG CAA ATT CAG CGG CTT GGC ATA AAG CAA GAC TCT GAA ATG AGG 1824
 556 I   L   L   A   G   K   Q   I   Q   R   L   G   I   K   Q   D   S   E   M   R   575

1825 AGT ATG CAA TCT GCA ACT GTG CGA TCT ACT CTT TTG AGA AAA TTG ATT TCT GTT ATT TCA 1884
 576 S   M   Q   S   A   T   V   R   S   T   L   L   R   K   L   I   S   V   I   S   595

1885 TCC CCT GTT GTG GTT GAC AAT GCC GGA AAA CTT CTC TCT GCC CTA AAT AAG GAA GCG GCT 1944
 596 S   P   V   V   V   D   N   A   G   K   L   L   S   A   L   N   K   E   A   A   615

1945 GTT CGA GGT GAC TTG CTC GAC ATA CTA ATC ACT TCC AGC GAC CAA TTT CCT GAG CTT GCT 2004
 616 V   R   G   D   L   L   D   I   L   I   T   S   S   D   Q   F   P   E   L   A   635

2005 GAA GCT CGC CAA GCA GTT TTA GTC ATC AGG GAA AAG CTG GAT TCC TCG ATA GCT TCA TTT 2064
 636 E   A   R   Q   A   V   L   V   I   R   E   K   L   D   S   S   I   A   S   F   655

2065 CGC AAG AAG CTC GCT ATT CGA AAT TTG GAA TTT CTT CAA GTG TCG GGG ATC ACA CAT TTG 2124
 656 R   K   K   L   A   I   R   N   L   E   F   L   Q   V   S   G   I   T   H   L   675

2125 ATA GAG CTG CCC GTT GAT TCC AAG GTC CCT ATG AAT TGG GTG AAA GTA AAT AGC ACC AAG 2184
 676 I   E   L   P   V   D   S   K   V   P   M   N   W   V   K   V   N   S   T   K   695

2185 AAG ACT ATT CGA TAT CAT CCC CCA GAA ATA GTA GCT GGC TTG GAT GAG CTA GCT CTA GCA 2244
 696 K   T   I   R   Y   H   P   P   E   I   V   A   G   L   D   E   L   A   L   A   715

2245 ACT GAA CAT CTT GCC ATT GTG AAC CGA GCT TCG TGG GAT AGT TTC CTC AAG AGT TTC AGT 2304
 716 T   E   H   L   A   I   V   N   R   A   S   W   D   S   F   L   K   S   F   S   735

2305 AGA TAC TAC ACA GAT TTT AAG GCT GCC GTT CAA GCT CTT GCT GCA CTG GAC TGT TTG CAC 2364
 736 R   Y   Y   T   D   F   K   A   A   V   Q   A   L   A   A   L   D   C   L   H   755

2365 TCC CTT TCA ACT CTA TCT AGA AAC AAG AAC TAT GTC CGT CCC GAG TTT GTG GAT GAC TGT 2424
 756 S   L   S   T   L   S   R   N   K   N   Y   V   R   P   E   F   V   D   D   C   775

2425 GAA CCA GTT GAG ATA AAC ATA CAG TCT GGT CGT CAT CCT GTA CTG GAG ACT ATA TTA CAA 2484
 776 E   P   V   E   I   N   I   Q   S   G   R   H   P   V   L   E   T   I   L   Q   795

2485 GAT AAC TTC GTC CCA AAT GAC ACA ATT TTG CAT GCA GAA GGG GAA TAT TGC CAA ATT ATC 2544
 796 D   N   F   V   P   N   D   T   I   L   H   A   E   G   E   Y   C   Q   I   I   815

2545 ACC GGA CCT AAC ATG GGA GGA AAG AGC TGC TAT ATC CGT CAA GTT GCT TTA ATT TCC ATA 2604
 816 T   G   P   N   M   G   G   K   S   C   Y   I   R   Q   V   A   L   I   S   I   835

2605 ATG GCT CAG GTT GGT TCC TTT GTA CCA GCG TCA TTC GCC AAG CTG CAC GTG CTT GAT GGT 2664
 836 M   A   Q   V   G   S   F   V   P   A   S   F   A   K   L   H   V   L   D   G   855

2665 GTT TTC ACT CGG ATG GGT GCT TCA GAC AGT ATC CAG CAT GGC AGA AGT ACC TTT CTA GAA 2724
 856 V   F   T   R   M   G   A   S   D   S   I   Q   H   G   R   S   T   F   L   E   875

2725 GAA TTA AGT GAA GCG TCA CAC ATA ATC AGA ACC TGT TCT TCT CGT TCG CTT GTT ATA TTA 2784
 876 E   L   S   E   A   S   H   I   I   R   T   C   S   S   R   S   L   V   I   L   895

2785 GAT GAG CTT GGA AGA GGC ACT AGC ACA CAC GAC GGT GTA GCC ATT GCC TAT GCA ACA TTA 2844
 896 D   E   L   G   R   G   T   S   T   H   D   G   V   A   I   A   Y   A   T   L   915

2845 CAG CAT CTC CTA GCA GAA AAG AGA TGT TTG GTT CTT TTT GTC ACG CAT TAC CCT GAA ATA 2904
 916 Q   H   L   L   A   E   K   R   C   L   V   L   F   V   T   H   Y   P   E   I   935

2905 GCT GAG ATC AGT AAC GGA TTC CCA GGT TCT GTT GGG ACA TAC CAT GTC TCG TAT CTG ACA 2964
 936 A   E   I   S   N   G   F   P   G   S   V   G   T   Y   H   V   S   Y   L   T   955

2965 TTG CAG AAG GAT AAA GGC AGT TAT GAT CAT GAT GAT GTG ACC TAC CTA TAT AAG CTT GTG 3024
 956 L   Q   K   D   K   G   S   Y   D   H   D   D   V   T   Y   L   Y   K   L   V   975

3025 CGT GGT CTT TGC AGC AGG AGC TTT GGT TTT AAG GTT GCT CAG CTT GCC CAG ATA CCT CCA 3084
 976 R   G   L   C   S   R   S   F   G   F   K   V   A   Q   L   A   Q   I   P   P   995

3085 TCA TGT ATA CGT CGA GCC ATT TCA ATG GCT GCA AAA TTG GAA GCT GAG GTA CGT GCA AGA 3144
 996 S   C   I   R   R   A   I   S   M   A   A   K   L   E   A   E   V   R   A   R  1015

3145 GAG AGA AAT ACA CGC ATG GGA GAA CCA GAA GGA CAT GAA GAA CCG AGA GGC GCA GAA GAA 3204
1016 E   R   N   T   R   M   G   E   P   E   G   H   E   E   P   R   G   A   E   E  1035

3205 TCT ATT TCG GCT CTA GGT GAC TTG TTT GCA GAC CTG AAA TTT GCT CTC TCT GAA GAG GAC 3264
1036 S   I   S   A   L   G   D   L   F   A   D   L   K   F   A   L   S   E   E   D  1055
```

Figure 4 (continued)

```
3265 CCT TGG AAA GCA TTC GAG TTT TTA AAG CAT GCT TGG AAG ATT GCT GGC AAA ATC AGA CTA  3324
1056 P   W   K   A   F   E   F   L   K   H   A   W   K   I   A   G   K   I   R   L    1075

3325 AAA CCA ACT TGT TCA TTT TGA TTTAATCTTAACATTATAGCAACTGCAAGGTCTTGATCATCTGTTAGTTGCG 3397
1076 K   P   T   C   S   F   *                                                        1082

3398 TACTAACTT ATG TGT ATT AGT ATA ACA AGA AAA GAG AAT TAG AGAG ATG GAT TCT AAT CCG    3458
   1           M   C   I   S   I   T   R   K   E   N   *       M   D   S   N   P     5

3459 GTG TTG CAG TAC ATC TTT TCT CCA CCC GCA TAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA       3522
   6 V   L   Q   Y   I   F   S   P   P   A   *                                       16
```

Figure 5

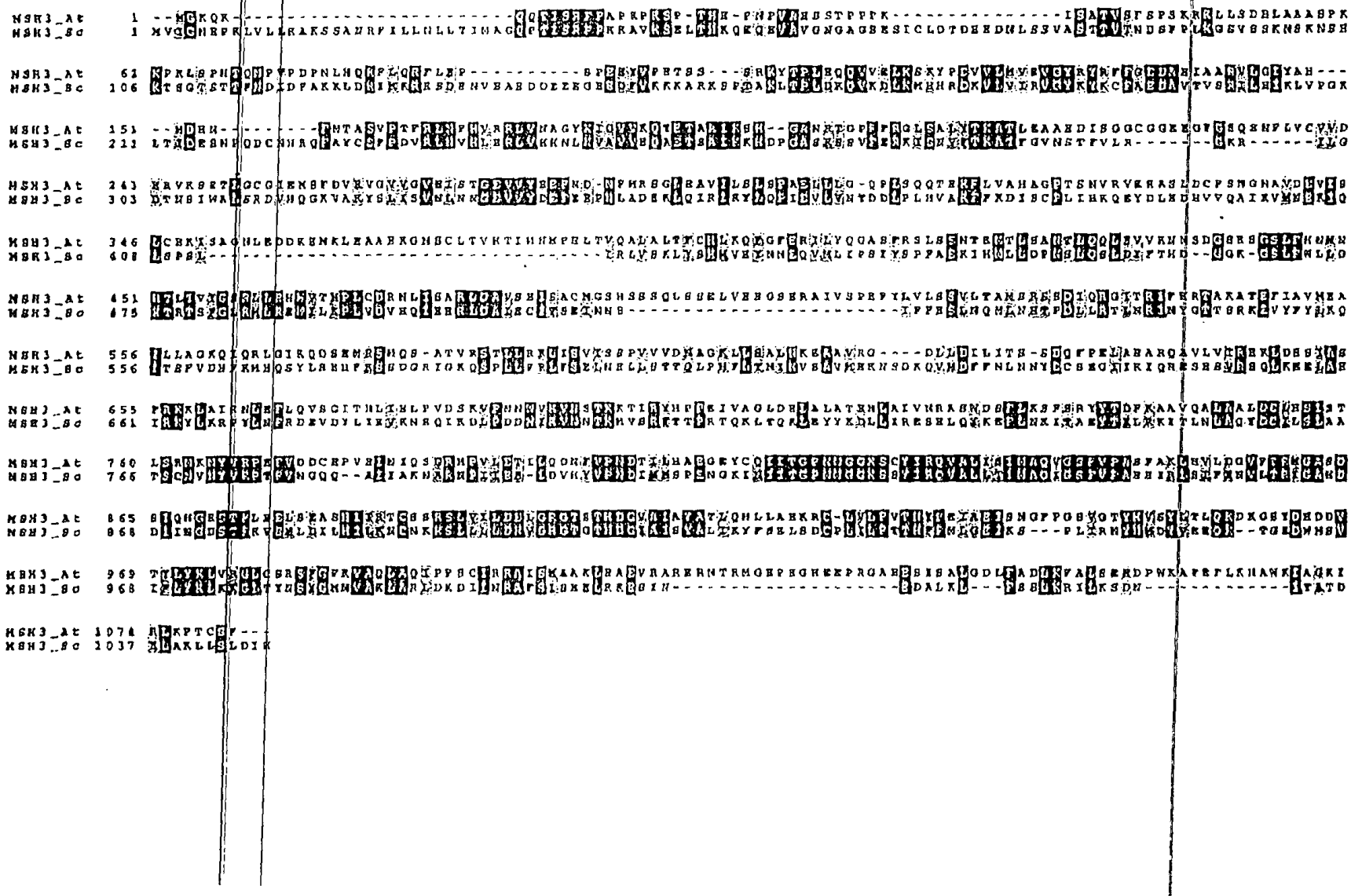
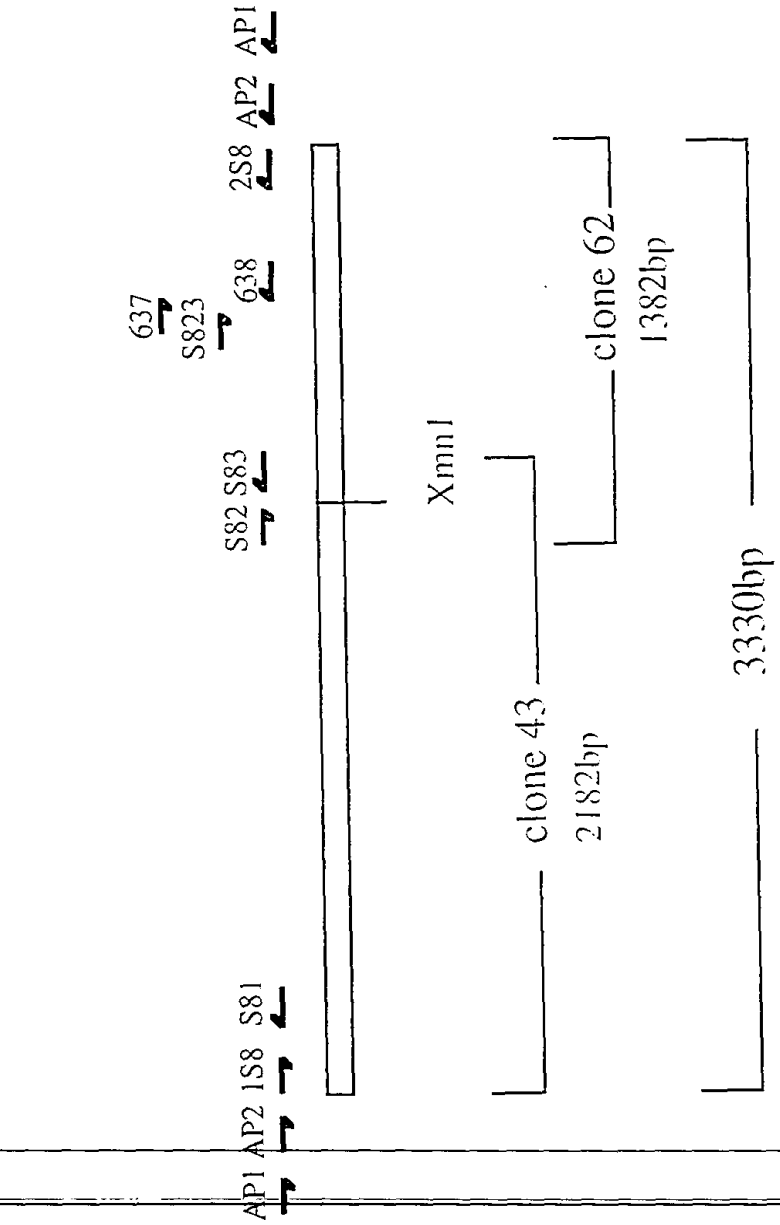


Figure 6



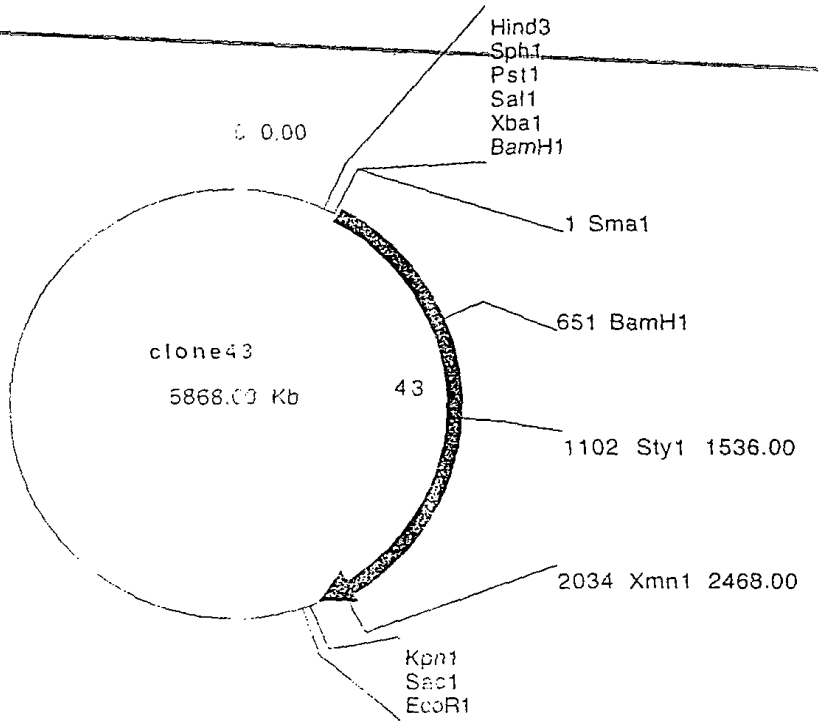


Figure 7

Comments/References: 43= 5' side of S8 (AtMSH6) 2182 bp in pUC18/Sma1

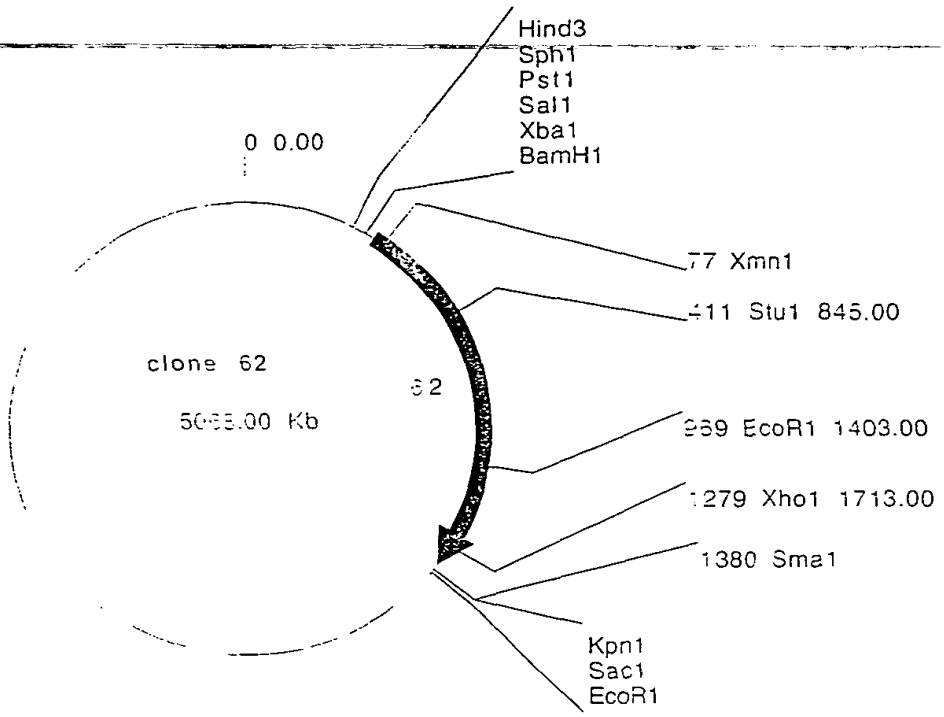


Figure 8

Comments/References: 62= 3' side of S8 (AtMSH6) 1379bp in pUC18/Sma1

Figure 9

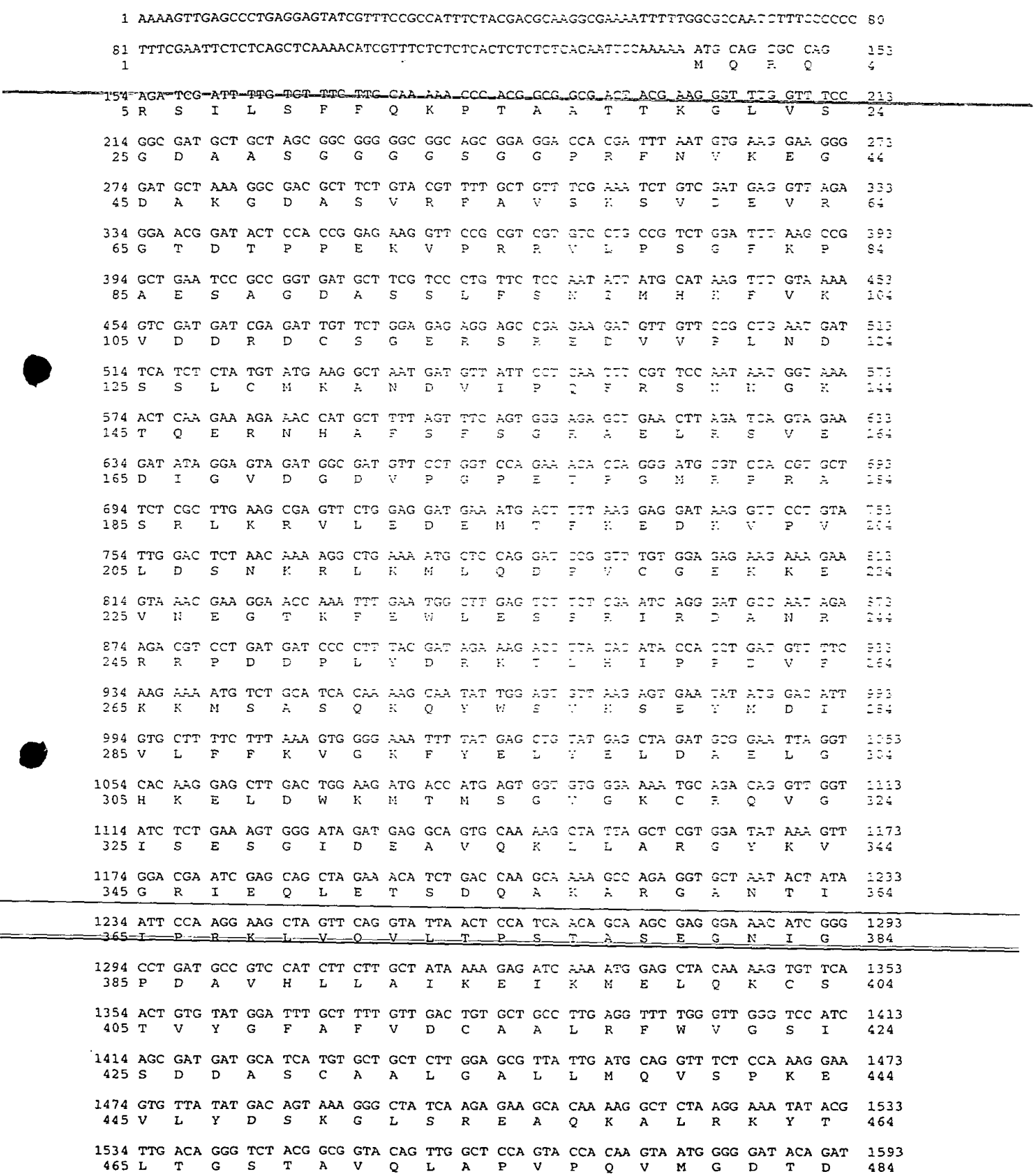

1 AAAAGTTGAGCCCTGAGGAGTATCGTTTCCGCCATTTCTACGACGCAAGGCGAAAATTTTTGGCGCCAATCTTTCCCCCC 80
81 TTTCGAATTCTCTCAGCTCAAAACATCGTTTCTCTCTCACTCTCTCTCACAATTCCAAAAA ATG CAG CGC CAG 153
1 M Q R Q 4
154 AGA TCG ATT TTG TCT TTC TTC CAA AAA CCC ACG GCG GCG ACT ACG AAG GGT TTG GTT TCC 213
5 R S I L S F F Q K P T A A T T K G L V S 24
214 GGC GAT GCT GCT AGC GGC GGG GGC GGC AGC GGA GGA CCA CGA TTT AAT GTG AAG GAA GGG 273
25 G D A A S G G G G S G G P R F N V K E G 44
274 GAT GCT AAA GGC GAC GCT TCT GTA CGT TTT GCT GTT TCG AAA TCT GTC GAT GAG GTT AGA 333
45 D A K G D A S V R F A V S K S V D E V R 64
334 GGA ACG GAT ACT CCA CCG GAG AAG GTT CCG CGT CGT GTC CTG CCG TCT GGA TTT AAG CCG 393
65 G T D T P P E K V P R R V L P S G F K P 84
394 GCT GAA TCC GCC GGT GAT GCT TCG TCC CTG TTC TCC AAT ATT ATG CAT AAG TTT GTA AAA 453
85 A E S A G D A S S L F S N I M H K F V K 104
454 GTC GAT GAT CGA GAT TGT TCT GGA GAG AGG AGC CGA GAA GAT GTT GTT CCG CTG AAT GAT 513
105 V D D R D C S G E R S R E D V V P L N D 124
514 TCA TCT CTA TGT ATG AAG GCT AAT GAT GTT ATT CCT CAA TTT CGT TCC AAT AAT GGT AAA 573
125 S S L C M K A N D V I P Q F R S N N G K 144
574 ACT CAA GAA AGA AAC CAT GCT TTT AGT TTC AGT GGG AGA GCT GAA CTT AGA TCA GTA GAA 633
145 T Q E R N H A F S F S G R A E L R S V E 164
634 GAT ATA GGA GTA GAT GGC GAT GTT CCT GGT CCA GAA ACA CCA GGG ATG CGT CCA CGT GCT 693
165 D I G V D G D V P G P E T P G M R P R A 184
694 TCT CGC TTG AAG CGA GTT CTG GAG GAT GAA ATG ACT TTT AAG GAG GAT AAG GTT CCT GTA 753
185 S R L K R V L E D E M T F K E D K V P V 204
754 TTG GAC TCT AAC AAA AGG CTG AAA ATG CTC CAG GAT CCG GTT TGT GGA GAG AAG AAA GAA 813
205 L D S N K R L K M L Q D P V C G E K K E 224
814 GTA AAC GAA GGA ACC AAA TTT GAA TGG CTT GAG TCT TCT CGA ATC AGG GAT GCC AAT AGA 873
225 V N E G T K F E W L E S S R I R D A N R 244
874 AGA CGT CCT GAT GAT CCC CTT TAC GAT AGA AAG ACC TTA GAC ATA CCA CCT GAT GTT TTC 933
245 R R P D D P L Y D R K T L H I P P D V F 264
934 AAG AAA ATG TCT GCA TCA CAA AAG CAA TAT TGG AGT GTT AAG AGT GAA TAT ATG GAC ATT 993
265 K K M S A S Q K Q Y W S V K S E Y M D I 284
994 GTG CTT TTC TTT AAA GTG GGG AAA TTT TAT GAG CTG TAT GAG CTA GAT GCG GAA TTA GGT 1053
285 V L F F K V G K F Y E L Y E L D A E L G 304
1054 CAC AAG GAG CTT GAC TGG AAG ATG ACC ATG AGT GGT GTG GGA AAA TGC AGA CAG GTT GGT 1113
305 H K E L D W K M T M S G V G K C R Q V G 324
1114 ATC TCT GAA AGT GGG ATA GAT GAG GCA GTG CAA AAG CTA TTA GCT CGT GGA TAT AAA GTT 1173
325 I S E S G I D E A V Q K L L A R G Y K V 344
1174 GGA CGA ATC GAG CAG CTA GAA ACA TCT GAC CAA GCA AAA GCC AGA GGT GCT AAT ACT ATA 1233
345 G R I E Q L E T S D Q A K A R G A N T I 364
1234 ATT CCA AGG AAG CTA GTT CAG GTA TTA ACT CCA TCA ACA GCA AGC GAG GGA AAC ATC GGG 1293
365 I P R K L V Q V L T P S T A S E G N I G 384
1294 CCT GAT GCC GTC CAT CTT CTT GCT ATA AAA GAG ATC AAA ATG GAG CTA CAA AAG TGT TCA 1353
385 P D A V H L L A I K E I K M E L Q K C S 404
1354 ACT GTG TAT GGA TTT GCT TTT GTT GAC TGT GCT GCC TTG AGG TTT TGG GTT GGG TCC ATC 1413
405 T V Y G F A F V D C A A L R F W V G S I 424
1414 AGC GAT GAT GCA TCA TGT GCT GCT CTT GGA GCG TTA TTG ATG CAG GTT TCT CCA AAG GAA 1473
425 S D D A S C A A L G A L L M Q V S P K E 444
1474 GTG TTA TAT GAC AGT AAA GGG CTA TCA AGA GAA GCA CAA AAG GCT CTA AGG AAA TAT ACG 1533
445 V L Y D S K G L S R E A Q K A L R K Y T 464
1534 TTG ACA GGG TCT ACG GCG GTA CAG TTG GCT CCA GTA CCA CAA GTA ATG GGG GAT ACA GAT 1593
465 L T G S T A V Q L A P V P Q V M G D T D 484

Figure 9 (continued)

```
1594 GCT GCT GGA GTT AGA AAT ATA ATA GAA TCT AAC GGA TAC TTT AAA GGT TCT TCT GAA TCA 1653
 485 A   A   G   V   R   N   I   I   E   S   N   G   Y   F   K   G   S   S   E   S  504

1654 TGG AAC TGT GCT GTT GAT GGT CTA AAT GAA TGT GAT GTT GCC CTT AGT GCT CTT GGA GAG 1713
 505 W   N   C   A   V   D   G   L   N   E   C   D   V   A   L   S   A   L   G   E  524

1714 CTA ATT AAT CAT CTG TCT AGG CTA AAG CTA GAA GAT GTA CTT AAG CAT GGG GAT ATT TTT 1773
 525 L   I   N   H   L   S   R   L   K   L   E   D   V   L   K   H   G   D   I   F  544

1774 CCA TAC CAA GTT TAC AGG GGT TGT CTC AGA ATT GAT GGC CAG ACG ATG GTA AAT CTT GAG 1833
 545 P   Y   Q   V   Y   R   G   C   L   R   I   D   G   Q   T   M   V   N   L   E  564

1834 ATA TTT AAC AAT AGC TGT GAT GGT GGT CCT TCA GGG ACC CTG TAC AAA TAT CTT GAT AAC 1893
 565 I   F   N   N   S   C   D   G   G   P   S   G   T   L   Y   K   Y   L   D   N  584

1894 TGT GTT AGT CCA ACT GGT AAG CGA CTC TTA AGG AAT TGG ATC TGC CAT CCA CTC AAA GAT 1953
 585 C   V   S   P   T   G   K   R   L   L   R   N   W   I   C   H   P   L   K   D  604

1954 GTA GAA AGC ATC AAT AAA CGG CTT GAT GTA GTT GAA GAA TTC ACG GCA AAC TCA GAA AGT 2013
 605 V   E   S   I   N   K   R   L   D   V   V   E   E   F   T   A   N   S   E   S  624

2014 ATG CAA ATC ACT GGC CAG TAT CTC CAC AAA CTT CCA GAC TTA GAA AGA CTG CTC GGA CGC 2073
 625 M   Q   I   T   G   Q   Y   L   H   K   L   P   D   L   E   R   L   L   G   R  644

2074 ATC AAG TCT AGC GTT CGA TCA TCA GCC TCT GTG TTG CCT GCT CTT CTG GGG AAA AAA GTG 2133
 645 I   K   S   S   V   R   S   S   A   S   V   L   P   A   L   L   G   K   K   V  664

2134 CTG AAA CAA CGA GTT AAA GCA TTT GGG CAA ATT GTG AAA GGG TTC AGA AGT GGA ATT GAT 2193
 665 L   K   Q   R   V   K   A   F   G   Q   I   V   K   G   F   R   S   G   I   D  684

2194 CTG TTG TTG GCT CTA CAG AAG GAA TCA AAT ATG ATG AGT TTG CTT TAT AAA CTC TGT AAA 2253
 685 L   L   L   A   L   Q   K   E   S   N   M   M   S   L   L   Y   K   L   C   K  704

2254 CTT CCT ATA TTA GTA GGA AAA AGC GGG CTA GAG TTA TTT CTT TCT CAA TTC GAA GCA GCC 2313
 705 L   P   I   L   V   G   K   S   G   L   E   L   F   L   S   Q   F   E   A   A  724

2314 ATA GAT AGC GAC TTT CCA AAT TAT CAG AAC CAA GAT GTG ACA GAT GAA AAC GCT GAA ACT 2373
 725 I   D   S   D   F   P   N   Y   Q   N   Q   D   V   T   D   E   N   A   E   T  744

2374 CTC ACA ATA CTT ATC GAA CTT TTT ATC GAA AGA GCA ACT CAA TGG TCT GAG GTC ATT CAC 2433
 745 L   T   I   L   I   E   L   F   I   E   R   A   T   Q   W   S   E   V   I   H  764

2434 ACC ATA AGC TGC CTA GAT GTC CTG AGA TCT TTT GCA ATC GCA GCA AGT CTC TCT GCT GGA 2493
 765 T   I   S   C   L   D   V   L   R   S   F   A   I   A   A   S   L   S   A   G  784

2494 AGC ATG GCC AGG CCT GTT ATT TTT CCC GAA TCA GAA GCT ACA GAT CAG AAT CAG AAA ACA 2553
 785 S   M   A   R   P   V   I   F   P   E   S   E   A   T   D   Q   N   Q   K   T  804

2554 AAA GGG CCA ATA CTT AAA ATC CAA GGA CTA TGG CAT CCA TTT GCA GTT GCA GCC GAT GGT 2613
 805 K   G   P   I   L   K   I   Q   G   L   W   H   P   F   A   V   A   A   D   G  824

2614 CAA TTG CCT GTT CCG AAT GAT ATA CTC CTT GGC GAG GCT AGA AGA AGC AGT GGC AGC ATT 2673
 825 Q   L   P   V   P   N   D   I   L   L   G   E   A   R   R   S   S   G   S   I  844

2674 CAT CCT CGG TCA TTG TTA CTG ACG GGA CCA AAC ATG GGC GGA AAA TCA ACT CTT CTT CGT 2733
 845 H   P   R   S   L   L   L   T   G   P   N   M   G   G   K   S   T   L   L   R  864

2734 GCA ACA TGT CTG GCC GTT ATC TTT GCC CAA CTT GGC TGC TAC GTG CCG TGT GAG TCT TGC 2793
 865 A   T   C   L   A   V   I   F   A   Q   L   G   C   Y   V   P   C   E   S   C  884

2794 GAA ATC TCC CTC GTG GAT ACT ATC TTC ACA AGG CTT GGC GCA TCT GAT AGA ATC ATG ACA 2853
 885 E   I   S   L   V   D   T   I   F   T   R   L   G   A   S   D   R   I   M   T  904

2854 GGA GAG AGT ACC TTT TTG GTA GAA TGC ACT GAG ACA GCG TCA GTT CTT CAG AAT GCA ACT 2913
 905 G   E   S   T   F   L   V   E   C   T   E   T   A   S   V   L   Q   N   A   T  924

2914 CAG GAT TCA CTA GTA ATC CTT GAC GAA CTG GGC AGA GGA ACT AGT ACT TTC GAT GGA TAC 2973
 925 Q   D   S   L   V   I   L   D   E   L   G   R   G   T   S   T   F   D   G   Y  944

2974 GCC ATT GCA TAC TCG GTT TTT CGT CAC CTG GTA GAG AAA GTT CAA TGT CGG ATG CTC TTT 3033
 945 A   I   A   Y   S   V   F   R   H   L   V   E   K   V   Q   C   R   M   L   F  964

3034 GCA ACA CAT TAC CAC CCT CTC ACC AAG GAA TTC GCG TCT CAC CCA CGT GTC ACC TCG AAA 3093
 965 A   T   H   Y   H   P   L   T   K   E   F   A   S   H   P   R   V   T   S   K  984

3094 CAC ATG GCT TGC GCA TTC AAA TCA AGA TCT GAT TAT CAA CCA CGT GGT TGT GAT CAA GAC 3153
 985 H   M   A   C   A   F   K   S   R   S   D   Y   Q   P   R   G   C   D   Q   D 1004

3154 CTA GTG TTC TTG TAC CGT TTA ACC GAG GGA GCT TGT CCT GAG AGC TAC GGA CTT CAA GTG 3213
1005 L   V   F   L   Y   R   L   T   E   G   A   C   P   E   S   Y   G   L   Q   V 1024

3214 GCA CTC ATG GCT GGA ATA CCA AAC CAA GTG GTT GAA ACA GCA TCA GGT GCT GCT CAA GCC 3273
1025 A   L   M   A   G   I   P   N   Q   V   V   E   T   A   S   G   A   A   Q   A 1044
```

Figure 9 (continued)

```
3274 ATG AAG AGA TCA ATT GGG GAA AAC TTC AAG TCA AGT GAG CTA AGA TCT GAG TTC TCA AGT  3333
1045 M   K   R   S   I   G   E   N   F   K   S   S   E   L   R   S   E   F   S   S    1064

3334 CTG CAT GAA GAC TGG CTC AAG TCA TTG GTG GGT ATT TCT CGA GTC GCC CAC AAC AAT GCC  3393
1065 L   H   E   D   W   L   K   S   L   V   G   I   S   R   V   A   H   N   N   A    1084

3394 CCC ATT GGC GAA GAT GAC TAC GAC ACT TTG TTT TGC TTA TGG CAT GAG ATC AAA TCC TCT  3453
1085 P   I   G   E   D   D   Y   D   T   L   F   C   L   W   H   E   I   K   S   S    1104

3454 TAC TGT GTT CCC AAA TAA  ATG GCT ATG ACA TAA CACTATCTGAAGCTCGTTAAGTCTTTTGCCTCTCT 3521
1105 Y   C   V   P   K   *    M   A   M   T   *                                       5

3522 G ATG TTT ATT CCT CTT AAA AAA TGC TTA TAT ATC AAA AAA TTG TTT CCT CGA TTA AAA    3579
   1   M   F   I   P   L   K   K   C   L   Y   I   K   K   L   F   P   R   L   K     19

3580 AAA AAA AAA AAA AAA AAA AAA AAA AAA                                             3606
  20 K   K   K   K   K   K   K   K   K                                               28
```

Figure 10

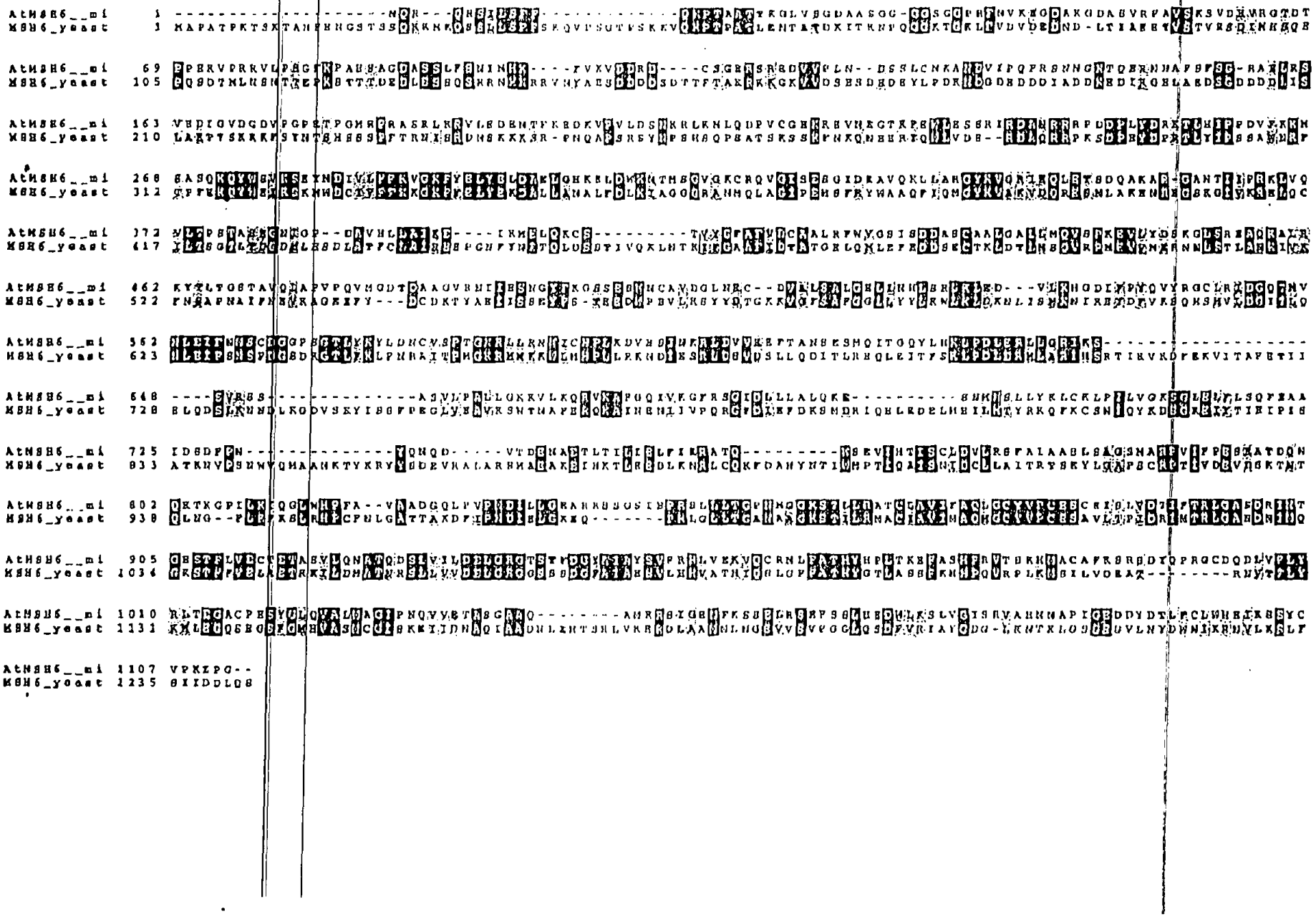


Figure 11

Complete genomic DNA sequence, 8062 bp, of *AtMSH6*

```
TTTTTTGGTTGCTAACAATAAAGGTATACGGTTTTATGTCATCAATATAA      50
CTATATATAAAAGAAATGAAAGATATATATTGTTTTTTCATTTATCAAAC     100
AAAACAACAAGACTTTTTTTTTTACTTTTTACATTGGTCAACAAAATACAA    150
GATAAACGACATCGTTTAATCATTTCCCAATTTTACCCCTAAGTTTAACA     200
CCTAGAACCTTCTCCATCTTCGCAAGCACAGCCTGATTAGGAACAGCTTT     250
ACCATTCTCATATTCCTGAACTACCTGAGTCCTCTCATTGATCTGTTTCG     300
CCAAATCCGCTTGTGACATCTTCTTCTCCAATCTCGCTTTCTGTATCATC     350
AACCTCACCTCTGCTTTCACACGATCCATCGCCGCAGGCTCTGTTTCTTC     400
TTCCAGCTTCTTCGTGTTAATCACCGGAACCGCCGTAGATTTCCCCTTTT     450
TGTTCGAACCGGCATCGAATTTCTTAACCGTTTGAACCGCGACACCGTTT     500
CTCAGAGCTGCGTTAACCGCTTTCGGATCGCGTAGGTCTTGGCTCTTTTG     550
TTTTGATTTGTGGAGAACTACTGGTTCCCAGTCTTGTGTTACTGCTCCTG     600
GGTATCTGCTCGGCATCGTCGATGAATTGAGAGAAAGGAACAACGCGAAA     650
ATTTTATTAATCTGAGTTTTGAAATTGAGAAACGATGAAGATGAAGAATG     700
TTGTTGAGAGGATTGTGATATTTATATATACGAAGATTGGTTTCTGGAGA     750
ATTCGATCATCTTTTTCTCCATTTTCGTCTCTGGAACGTTCTTAGAGATG     800
ATTGACGACGTGTCATTATCTGATTTGCAGTTAACCAATGCTTTTTGGGT     850
TGGATTCGTGGTACACCATATTATCCGATTTGGCTCAATGGTTTTATATA     900
AATTTGGTTTTCGGTTCGGTTATGAGTTATCATTAAAATTAAGCTAACCA     950
AAAATTTTCGTAAAATTTATTTCGGTTTCAATTCGGATCCCTTACTTCCA    1000
GAACCGAATTATTCGAAACCGGGGTTAGCCGAACCGAATACCAATGCCTG    1050
ATTGACTCGTTGGCTAGAAAGATCCAACGGTATACAATAATAGAACATAA    1100
ATCGGACGGTCATCAAAGCCTCAAAGAGTGAACAGTCAACAAAAAAAGTT    1150
GAGCCCTGAGGAGTATCGTTTCCGCCATTTCTACGACGCAAGGCGAAAAT    1200
TTTTGGCGCCAATCTTTCCCCCCTTTCGAATTCTCTCAGCTCAAAACATC    1250
GTTTCTCTCTCACTCTCTCTCACAATTCCAAAAAATGCAGCGCCAGAGAT    1300
CGATTTTGTCTTTCTTCCAAAAACCCACGGCGGCGACTACGAAGGGTTTG    1350
GTTTCCGGCGATGCTGCTAGCGGCGGGGGCGGCAGCGGAGACCACGATTT    1400
AATGTGAAGGAAGGGGATGCTAAAGGCGACGCTTCTGTACGTTTTGCTGT    1450
TTCGAAATCTGTCGATGAGGTTAGAGGAACGGATACTCCACCGGAGAAGG    1500
TTCCGCGTCGTGTCCTGCCGTCTGGATTTAAGCCGGCTGAATCCGCCGGT    1550
GATGCTTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA    1600
TGATCGAGATTGTTCTGGAGAGAGGTACTAATCTTCGATTCTCTTAATTT    1650
TGTTATCTTTAGCTGGAAGAAGAAGATTCGTGTAATTTGTTGTATTCGTT    1700
GGAGAGATTCTGATTACTGCATTGGATCGTTGTTTACAAATTTTCAGGAG    1750
CCGAGAAGATGTTGTTCCGCTGAATGATTCATCTCTATGTATGAAGGCTA    1800
ATGATGTTATTCCTCAATTTCGTTCCAATAATGGTAAAACTCAAGAAAGA    1850
AACCATGCTTTTAGTTTCAGTGGGAGAGCTGAACTTAGATCAGTAGAAGA    1900
TATAGGAGTAGATGGCGATGTTCCTGGTCCAGAAACACCAGGGATGCGTC    1950
CACGTGCTTCTCGCTTGAAGCGAGTTCTGGAGGATGAAATGACTTTTAAG    2000
GAGGATAAGGTTCCTGTATTGGACTCTAACAAAAGGCTGAAAATGCTCCA    2050
GGATCCGGTTTGTGGAGAGAAGAAAGAAGTAAACGAAGGAACCAAATTTG    2100
AATGGCTTGAGTCTTCTCGAATCAGGGATGCCAATAGAAGACGTCCTGAT    2150
GATCCCCTTTACGATAGAAAGACCTTACACATACCACCTGATGTTTTCAA    2200
GAAAATGTCTGCATCACAAAAGCAATATTGGAGTGTTAAGAGTGAATATA    2250
TGGACATTGTGCTTTTCTTTAAAGTGGTTAGTAACTATTAATCTAGTGTT    2300
CAATCCATTTCCTCAATGTGATTTGTTCACTTACATCTGTTTACGTTATG    2350
CTCTTCTCAGGGGAAATTTTATGAGCTGTATGAGCTAGATGCGGAATTAG    2400
GTCACAAGGAGCTTGACTGGAAGATGACCATGAGTGGTGTGGGAAAATGC    2450
AGACAGGTAAATTAGTTGAAACAACTGGCCTGCTTGAATTATTGTGTCTA    2500
```

Figure 11 (Continued)

```
TAAATTTTGACACCACCTTTTGTTTCAGGTTGGTATCTCTGAAAGTGGGA      2550
TAGATGAGGCAGTGCAAAAGCTATTAGCTCGTGGGTAAGGGAACCATCAT      2600
ACTTTATGGAATTCGTTTACTGCTACTTCGGCTAGGATTTAAGAAATGGA      2650
AATCACTTCAAGCATCATTAGTTAGGATCCTGAGAACTCAGGATGTTTTC      2700
TTATTCGTTATATAATAAGTCTTTTCATCAAGGAGTAACAAACAAAACTT      2750
GCACAATATTTGTGTGCTCACTGGCAAGGCATATATACCCAGCTAACCTT      2800
TGCTAGTTCACTGTAGTAACAGTTACGGATAATATATGTTTACTTGTATG      2850
TGGTACCCTCATTTTGTCTCTCATGGAGGCTTTCAAGCCTTGTGTTGAAA      2900
CTGGATAGTTACATATGCTTCCAACAGAAACTAGCATGCAGATTCATATG      2950
CTTTCCTATTCTACTAATTATGTATTGACACACTCGTTGTTTCTTTTGAA      3000
AGATATAAAGTTGGACGAATCGAGCAGCTAGAAACATCTGACCAAGCAAA      3050
AGCCAGAGGTGCTAATACTGTAAGTTTTCTTGGATAGGTCAAGGAGAGTG      3100
TTGCAGACTGTTTTTGATCATTTCTTTTTCTGTACATTACTTTCATGCTG      3150
TAATTAACTCAATGGCTATTCTGGTCTGATTATCAGATAATTCCAAGGAA      3200
GCTAGTTCAGGTATTAACTCCATCAACAGCAAGCGAGGGAAACATCGGGC      3250
CTGATGCCGTCCATCTTCTTGCTATAAAAGAGGTTTGTTATTTACTTATT      3300
TATCTTATCATGTTCAGTTCATCCAAGTCCTGAAAAATTACACTCTTCTT      3350
TACCAATCTTCCATCAAGCTGTGTAAAGGATTTGGAATTAGAAAATCATT      3400
ATTTGATGCTTTGTTTTATATGCAAGAGGTTCCCTTGAAAAGATCTGTTT      3450
AAGATTCTTTGCACTTGAAAAATTCAATCTTTTTAAGTGAATCCCCTACT      3500
TTCTTACAATGATCATAGTCTGCAATTGCATGTCAAGTAATATCATTCCT      3550
TGTTACTGCATCCCCCTCTTTCTTAATGACCATTGTCTATGTTGTGTTTG      3600
TCTCGTGTGCTGGAGAAAATGATAGCTGATCCAAGCTGTACATTATCATG      3650
ATTAAGTAGCTGCTCAGGAATTGCCTTTGGTTACATTGCCTAATGGTTTG      3700
ATGTCAATTTTTCTTCTGAATCTTTATTTTAGATCAAAATGGAGCTACAA      3750
AAGTGTTCAACTGTGTATGGATTTGCTTTTGTTGACTGTGCTGCCTTGAG      3800
GTTTTGGGTTGGGTCCATCAGCGATGATGCATCATGTGCTGCTCTTGGAG      3850
CGTTATTGATGCAGGTAAGCAAGTGTATTCTGTATCTTATGTGTACCATG      3900
TGACTTCCTGTGCATATATTTGGGTTGCAGGAACTAATTCTGAATCACCA      3950
TTTGGTATGTTTTTTCCAGGTTTCTCCAAAGGAAGTGTTATATGACAGTA      4000
AAGGTAAACTGCTTGTATCGCCAGTTGTTTTGTTAAACAGAATTTAAGGT      4050
AAATGACACTGGTTAATTTAAAGTGCATACATGTTGAAATATTGCAGGGC      4100
TATCAAGAGAAGCACAAAAGGCTCTAAGGAAATATACGTTGACAGGTACC      4150
ATTTCAGTAGGCAAGCTAACTGACAATTTAACCGCTCACCGAATGATAGG      4200
TCTCTTAAACATTGCTAATGTAGATGATGTTTATGTTTCAATCTAATAGG      4250
GTCTACGGCGGTACAGTTGGCTCCAGTACCACAAGTAATGGGGGATACAG      4300
ATGCTGCTGGAGTTAGAAATATAATAGAATCTAACGGATACTTTAAAGGT      4350
TCTTCTGAATCATGGAACTGTGCTGTTGATGGTCTAAATGAATGTGATGT      4400
TGCCCTTAGTGCTCTTGGAGAGCTAATTAATCATCTGTCTAGGCTAAAGG      4450
TGTGTTGGCTTGTTTAGTTTTTGCTTTTCACAAATTAAGCAAAGGAACTT      4500
TTCATAACTTACAGTTTCTATCTACTTGCAGCTAGAAGATGTACTTAAGC      4550
ATGGGGATATTTTTCCATACCAAGTTTACAGGGGTTGTCTCAGAATTGAT      4600
GGCCAGACGATGGTAAATCTTGAGATATTTAACAATAGCTGTGATGGTGG      4650
TCCTTCAGGCAAGTGCATATTTCTTTTTTGATAACTTCAACTAGAGGGCA      4700
GACATAGAAGGAAAAATTCTAATACTTCGTACGGATCTCCAGTAAGTAAT      4750
AGCCGATTTTTGTTTACCTATGTAGGGACCTTGTACAAATATCTTGATAA      4800
CTGTGTTAGTCCAACTGGTAAGCGACTCTTAAGGAATTGGATCTGCCATC      4850
CACTCAAAGATGTAGAAAGCATCAATAAACGGCTTGATGTAGTTGAAGAA      4900
TTCACGGCAAACTCAGAAAGTATGCAAATCACTGGCCAGTATCTCCACAA      4950
ACTTCCAGACTTAGAAAGACTGCTCGGACGCATCAAGTCTAGCGTTCGAT      5000
CATCAGCCTCTGTGTTGCCTGCTCTTCTGGGGAAAAAAGTGCTGAAACAA      5050
CGAGTAAGTATCAATCACAAGTTTTCTGAGTAATGCCTTCCATGAGTAGT      5100
ATAGGACTAAAACATTACGGGTCTAGCTAAAGACTGTTCTCCTTCTTTTG      5150
```

Figure 11 (Continued)

```
CAATGTCTGGTTATTCATTACATTTCTCTTAACTTATTGCATTGCAGGTT      5200
AAAGCATTTGGGCAAATTGTGAAAGGGTTCAGAAGTGGAATTGATCTGTT      5250
GTTGGCTCTACAGAAGGAATCAAATATGATGAGTTTGCTTTATAAACTCT      5300
GTAAACTTCCTATATTAGTAGGAAAAAGCGGGCTAGAGTTATTTCTTTCT      5350
CAATTCGAAGCAGCCATAGATAGCGACTTTCCAAATTATCAGGTGCCCAT      5400
CTATCTTTCATACTTTACAACAAAATGTCTGTCACTACTCAAAGCAATGC      5450
ATATGGCTTAGATCTCAACTCACACCCCGAGGATCCTAAAGGGATTTGCT      5500
TTTTATTCCTAATGTTTTTGGATGGTTTGATTTATTTCTAACTTGAACTT      5550
ATTAATCTTGTACCAGAACCAAGATGTGACAGATGAAAACGCTGAAACTC      5600
TCACAATACTTATCGAACTTTTTATCGAAAGAGCAACTCAATGGTCTGAG      5650
GTCATTCACACCATAAGCTGCCTAGATGTCCTGAGATCTTTTGCAATCGC      5700
AGCAAGTCTCTCTGCTGGAAGCATGGCCAGGCCTGTTATTTTTCCCGAAT      5750
CAGAAGCTACAGATCAGAATCAGAAAACAAAAGGGCCAATACTTAAAATC      5800
CAAGGACTATGGCATCCATTTGCAGTTGCAGCCGATGGTCAATTGCCTGT      5850
TCCGAATGATATACTCCTTGGCGAGGCTAGAAGAAGCAGTGGCAGCATTC      5900
ATCCTCGGTCATTGTTACTGACGGGACCAAACATGGGCGGAAAATCAACT      5950
CTTCTTCGTGCAACATGTCTGGCCGTTATCTTTGCCCAAGTTTGTATACT      6000
CGTTAGATAATTACTCTATTCTTTGCAATCAGTTCTTCAACATGAATAAT      6050
AAATTCTGTTTTCTGTCTGCAGCTTGGCTGCTACGTGCCGTGTGAGTCTT      6100
GCGAAATCTCCCTCGTGGATACTATCTTCACAAGGCTTGGCGCATCTGAT      6150
AGAATCATGACAGGAGAGAGTAAGTTTTGTTCTCAAAATACCAATTCCTC      6200
GAACTATTTACTCAGATTTTGTCTGATTGGACAAGGTGGTTTTGCTTTTT      6250
TTTAGGTACCTTTTTGGTAGAATGCACTGAGACAGCGTCAGTTCTTCAGA      6300
ATGCAACTCAGGATTCACTAGTAATCCTTGACGAACTGGGCAGAGGAACT      6350
AGTACTTTCGATGGATACGCCATTGCATACTCGGTAACCTGCTCTTCTCC      6400
TTCAACTTATACTTGTTGATCAACAAAAACATGCAATTCATTTTGCTGAA      6450
ACTTATTGATTTATATCAGGTTTTTCGTCACCTGGTAGAGAAAGTTCAAT      6500
GTCGGATGCTCTTTGCAACACATTACCACCCTCTCACCAAGGAATTCGCG      6550
TCTCACCCACGTGTCACCTCGAAACACATGGCTTGCGCATTCAAATCAAG      6600
ATCTGATTATCAACCACGTGGTTGTGATCAAGACCTAGTGTTCTTGTACC      6650
GTTTAACCGAGGGAGCTTGTCCTGAGAGCTACGGACTTCAAGTGGCACTC      6700
ATGGCTGGAATACCAAACCAAGTGGTTGAAACAGCATCAGGTGCTGCTCA      6750
AGCCATGAAGAGATCAATTGGGGAAAACTTCAAGTCAAGTGAGCTAAGAT      6800
CTGAGTTCTCAAGTCTGCATGAAGACTGGCTCAAGTCATTGGTGGGTATT      6850
TCTCGAGTCGCCCACAACAATGCCCCCATTGGCGAAGATGACTACGACAC      6900
TTTGTTTTGCTTATGGCATGAGATCAAATCCTCTTACTGTGTTCCCAAAT      6950
AAATGGCTATGACATAACACTATCTGAAGCTCGTTAAGTCTTTTGCTTCT      7000
CTGATGTTTATTCCTCTTAAAAAATGCTTATATATCAAAAAATTGTTTCC      7050
TCGATTATAACAAGATTATATATGTATCTGTCGGTTTAGCTATGGTATAT      7100
AATATATGTATGTTCATGAGATTGGTCAAGAGAAATACTCACAAACAGTA      7150
TATTAAGAAGGAAATATGTTTATGCATTAATTTAAGTTTCAAGATAAACT      7200
GCAAATAACCTCGACTAAAGTTGCAAAGACCAAACACAAATTACAAAACT      7250
TATAAGACTTAAGTTCTGAATTCCCTAAAACCAAAAAAAAAAACAGAACA      7300
TATTTTGTTGCATCTACAAACAACACAAACCTACATAGTTTATAACTTAC      7350
TCATCACTGAGATTAACATCAGAATCATTCTCGATTTCTTCATCTTCACT      7400
CTCATCATCATCACCACCACCATGATGATTCTCCTCCTCTTCACGTAACC      7450
TAGCAATCTCACTCTGAGCTCTATCAACAATCTGCTTCTTCTGCAACTCC      7500
AAATCTCTCTGAAAATCAGCTCTCATCTTCTCCAACTCCTTCATTTGCTC      7550
TTTCTTACTCTTCTCCATCTTCTCATAAACCTTCCCAAACCTCTCAACAG      7600
AATCCGCCAACATCTTATACGAAGCAGCGTCATTAACCTTCTTCCTCTCG      7650
TACTCAACCTCATCATCCTCATCCTCCTCCTCTTCAGAATCACCAGGACT      7700
ATCCATCATCTCATCAAACCCATTAGACTTATCTAAATAAACCTTAGTGT      7750
TCATAAACACAAACTCACCTGAATCAACACCACAAGCTAAACCTAAATCC      7800
```

Figure 11 (Continued)

```
GACTTGGGCGAAACACAAAGCAACATATCCAACTTATTGAAAAACGACCA      7850
TTTACTTGAACCTAAACCTGATTTCTCAACCTTAATCTTCTCTTTTCTAT      7900
ACTTCCTCTTCAAGTCATCAATCATTCTCCTACATTGCGTCTCAGATTTC      7950
TCCATCCTTAGCTCCTCACTCACTTTCTCAGCTACTTCATTCCAATCCTC      8000
GTTCCTCAAACTCCTTCTACCCAATTGCAAAAACCTATCTCCCCAAACTT      8050
CAAGCAACACAA                                            8062
```

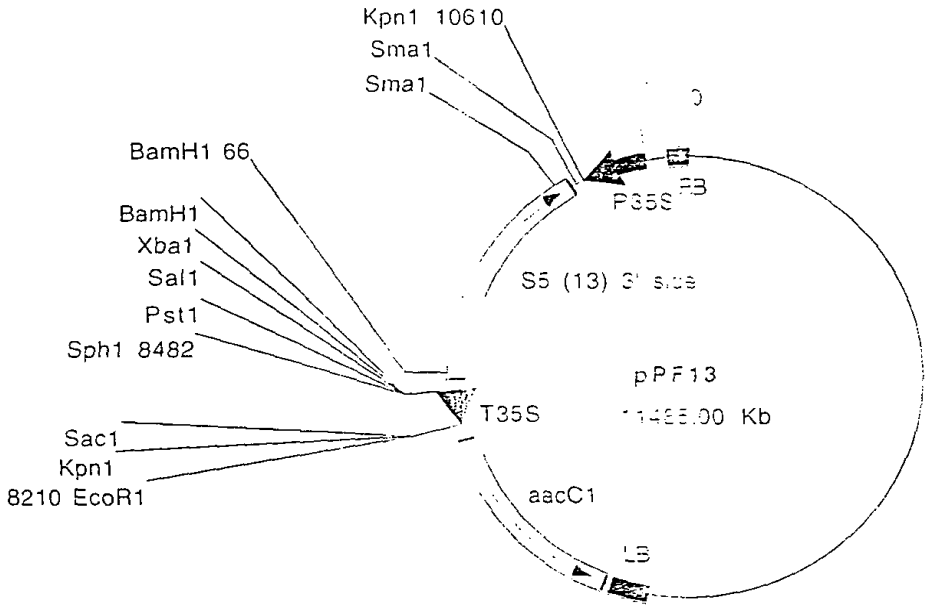


Figure 12

Comments/References: S5 3' side antisense : 13 from pUC18/13 Sal1/Sst1/T4 into 85/35S BamH1/T4 in Agrobacterium LBA4404

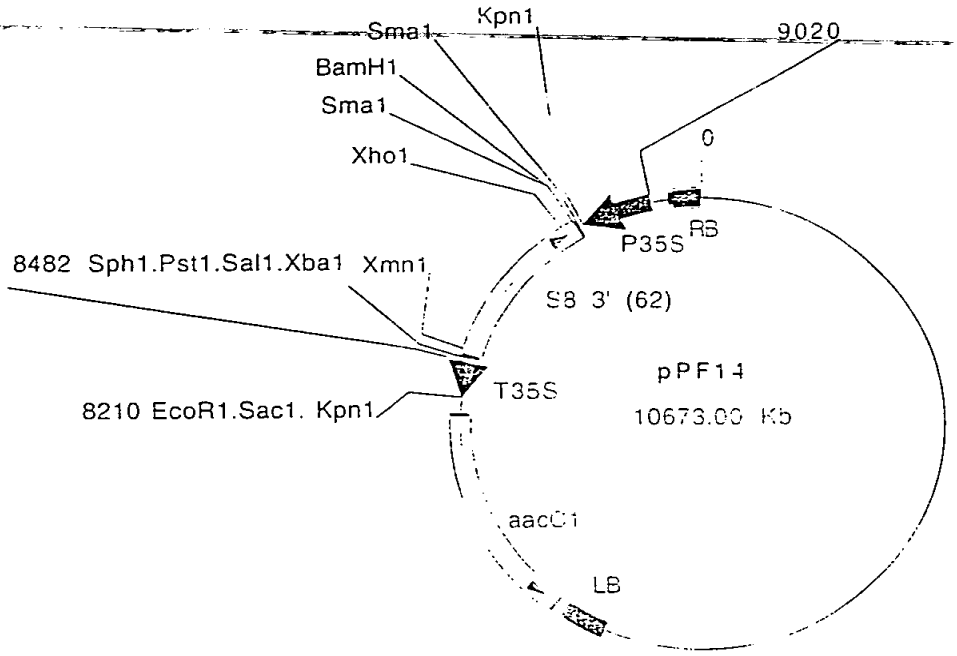


Figure 13

Comments/References: S8 3' side antisens : 62 Sal1/Sst1/T4 into 85 35S/BamH1/T4

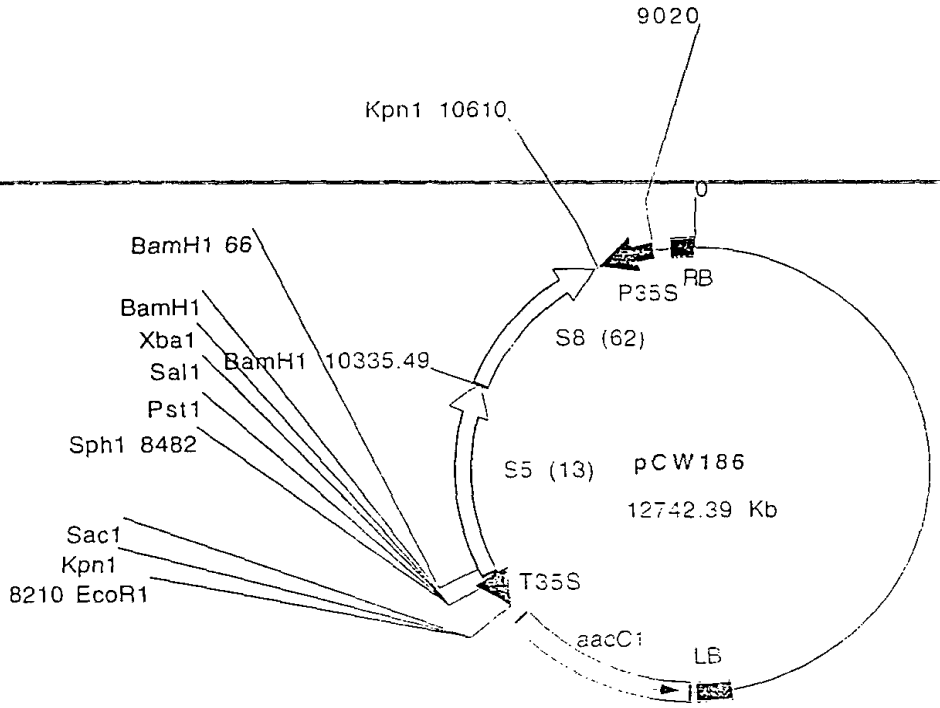


Figure 14

Comments/References: 85 35S S8 S5 antisense : S8 3' side (62) Sal1/Sst1/T4 into pPF13 (85 35S S5 3' side antisens)/Sma1. in LBA4404

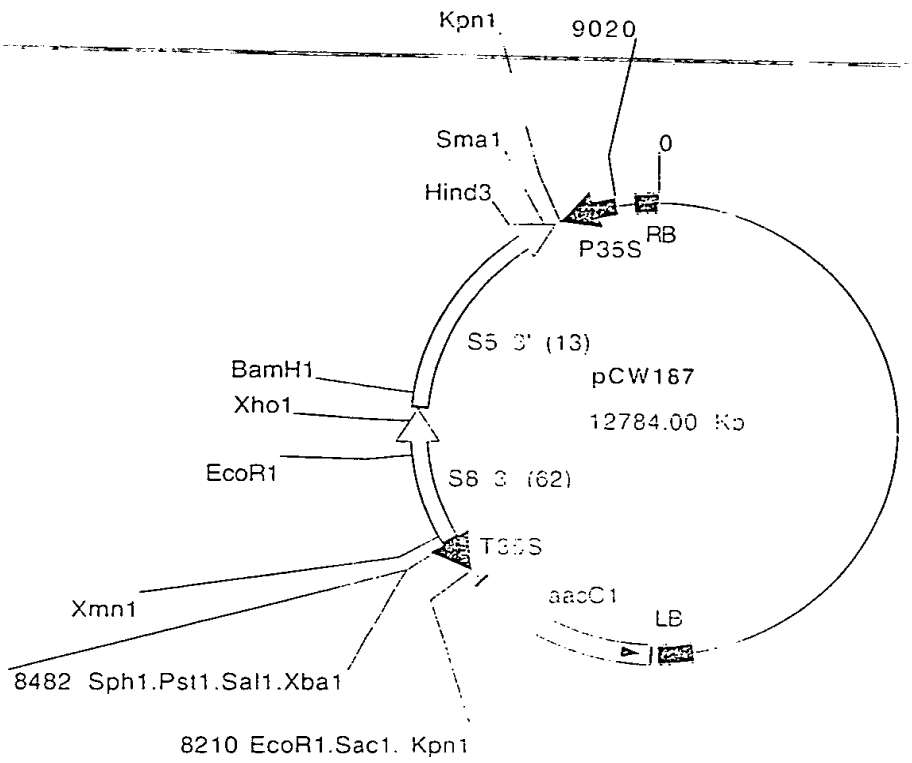
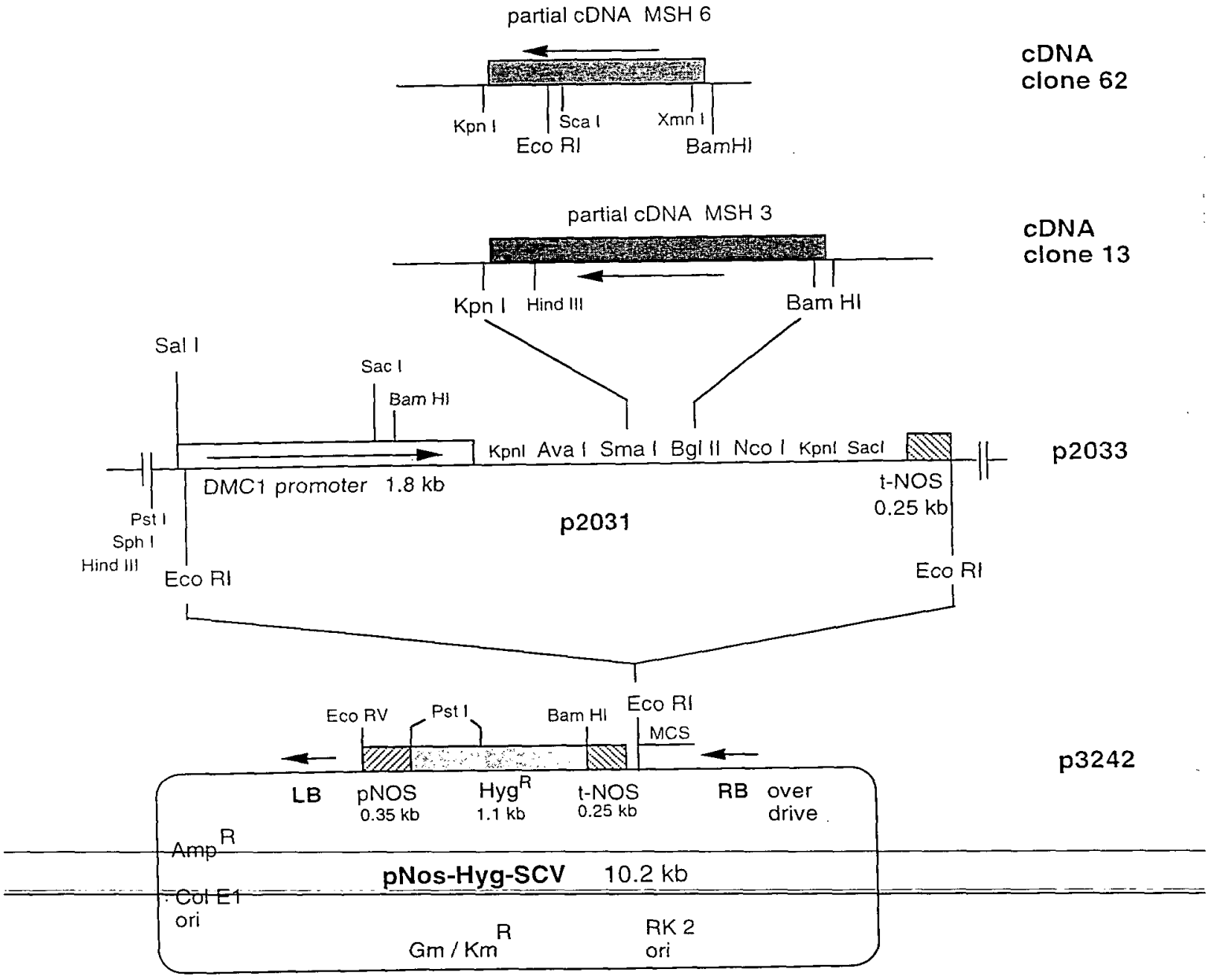


Figure 15

Comments/References: 85 35S S5.S8 antisens (D) : S5 3' side (13) Sal1/Sst1/T4 into pPF14 (85 35S S8 3' side antisense). in LBA4404

Figure 16


partial cDNA MSH 6
cDNA clone 62
Kpn I
Sca I
Xmn I
Eco RI
BamHI
partial cDNA MSH 3
cDNA clone 13
Kpn I
Hind III
Bam HI
Sal I
Sac I
Bam HI
KpnI Ava I Sma I Bgl II Nco I KpnI SacI
p2033
DMC1 promoter 1.8 kb
t-NOS 0.25 kb
p2031
Pst I
Sph I
Hind III
Eco RI
Eco RI
Eco RV
Pst I
Bam HI
Eco RI
MCS
p3242
LB
pNOS 0.35 kb
Hyg^R 1.1 kb
t-NOS 0.25 kb
RB
over drive
Amp^R
pNos-Hyg-SCV 10.2 kb
Col E1 ori
Gm / Km^R
RK 2 ori